ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ

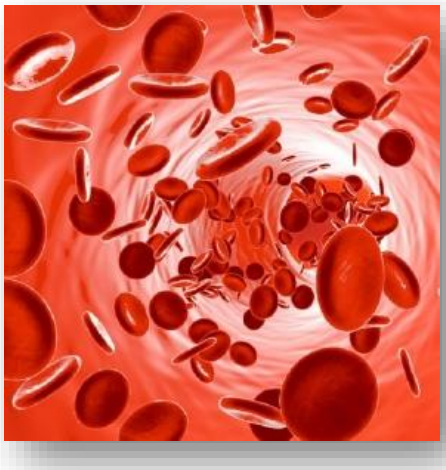
ΙΑΤΡΙΚΗ ΣΧΟΛΗ

Μεταπτυχιακή Διπλωματική Εργασία

Παραγωγή και μελέτη μονοκλωνικών αντισωμάτων από ασθενείς με νεοπλάσματα από ώριμα Β λεμφοκύτταρα

Αικατερίνη Καραγιάννη

Υπότροφος του Ιδρύματος Ωνάση



ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΕΡΕΥΝΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΗΣ ΑΝΑΠΤΥΞΗΣ

ΙΝΣΤΙΤΟΥΤΟ ΕΦΑΡΜΟΣΜΕΝΩΝ ΒΙΟΕΠΙΣΤΗΜΩΝ

Ηράκλειο, 2018

UNIVERSITY OF CRETE

SCHOOL OF HEALTH SCIENCES

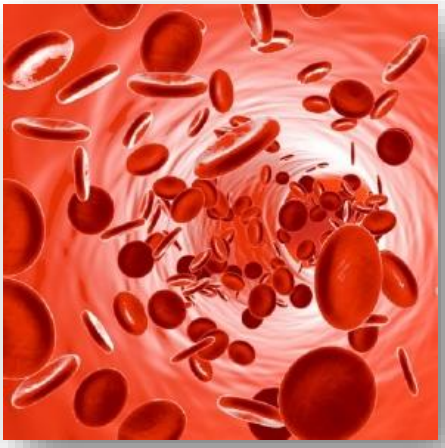
FACULTY OF MEDICINE

Postgraduate diploma thesis

Production and study of monoclonal antibodies from patients With Mature B cell Neoplasms

Aikaterini Karagianni

Alexander S. Onassis, Public Benefit Foundation Scholar



CENTRE FOR RESEARCH AND TECHNOLOGY HELLAS

INSTITUTE OF APPLIED BIOSCIENCES

Heraklion, 2018

Contents

Prologue

The B cell receptor immunoglobulin (BCR IG) is important for the ontogeny and evolution of malignancies of mature B cells (B cell non-Hodgkin lymphomas, B-NHLs), mediating critical interactions with the microenvironment, including recognition of and activation by antigens. Therefore, it is relevant to examine the antigenic specificity of BCR IGs in B-NHLs in an attempt to unravel the nature of the antigen(s) involved in the selection and expansion of the malignant clones.

Malignant B cells have a limited growing capacity *in vitro* and, at least for certain B-NHLs (most notably chronic lymphocytic leukemia, CLL) express low levels of surface IG. In order to overcome these problems, we used recombinant DNA technology for producing monoclonal antibodies from patients with CLL and splenic marinal zone lymphoma (SMZL) and then examine their antigenic reactivity profile.

The present study was conducted at the Institute of Applied Biosciences at the Centre for Research and Technology Hellas (INAB | CERTH), Thessaloniki, Greece. I thank Dr Kostas Stamatopoulos, Director of the Institute, for giving me the opportunity to conduct my thesis at INAB. I would also like to express my sincere gratitude to the postdoctoral researcher Maria Gounari, my "daily" supervisor, for her priceless assistance, guidance and patience throughout this year.

This thesis would not have been possible without the time and support of many people in the Institute: Nikos Papakonstantinou, Maria Koutroumani, Stamatia Laidou, Marina Gerousi, Konstantia Kotta, Spyros Petrakis, Andreadeli Angeliki, Tsiolas Giorgos, Maria Karypidou, Stauroula Ntoufa, Antonios Makris, Giorgos Gavriilidis, Giannis Gkekas, Despoina Binou and Sofia Papanikolaou. I am thankful for the inspiration and learning experience I gained from working with them and the precious moments we spent together in the lab.

I would not be who and where I am now without the love and endless support of my family. The constant encouragement I have gotten over the years gave me the strength to continue despite the difficulties.

Last but not least, I express my profound gratitude to Alexander S. Onassis Public Benefit Foundation for their support to this thesis.

Περίληψη

Η μελέτη των ανοσοσφαιρινών των Β κυτταρικών υποδοχέων σε μοριακό επίπεδο έχει συμβάλει αποφασιστικά στην κατανόηση της μοριακής βάσης κακοηθειών ώριμων Β λεμφοκυττάρων (B cell non-Hodgkin lymphomas, B-NHLs) και θεμελίωσε την ιδέα της αντιγονικής επιλογής στην ανάπτυξη και εξέλιξη της νόσου. Ωστόσο παρόλη την πρόοδο σε αυτό τον τομέα, δεν έχει αποσαφηνιστεί πλήρως ποια αντιγόνα είναι πιθανό να επιλέγουν ή και να διεγείρουν συνεχώς το νεοπλασματικό κλώνο. Για αυτό το λόγο, ανεξάρτητες μελέτες προσπαθούν να αποκαλύψουν την ακριβή φύση των αντιγόνων που εμπλέκονται στην παθογένεια των B-NHLs.

Στην παρούσα μελέτη επικεντρωθήκαμε σε 2 τύπους B-NHLs, τη Χρόνια Λεμφοκυτταρική Λευχαιμία (ΧΛΛ) και το Σπληνικό Λέμφωμα Οριακής Ζώνης (ΣΛΟΖ). Κλωνοτυπικές ανοσοσφαιρίνες Β κυτταρικών υποδοχέων (ΒΚΥ) παράχθηκαν ως ανασυνδυασμένα IgM μονοκλωνικά αντισώματα από 8 ασθενείς με ΣΛΟΖ και 34 ασθενείς με ΧΛΛ και ελέγχθηκαν ως προς την αντιγονική ειδικότητά τους με ανοσοδοκιμασίες ELISA, κυτταρομετρία ροής και ανοσοστύπωμα κατά Western.

Επιβεβαιώσαμε ότι περιπτώσεις ασθενών με ΧΛΛ που ανήκουν στο ίδιο στερεότυπο υποσύνολο, δηλαδή εκφράζουν ομόλογους Β κυτταρικούς υποδοχείς (ΒκΥ), παρουσίασαν παρόμοιο προφίλ αντιγονικής αναγνώρισης. Ενδιαφέρον ήταν επίσης το γεγονός πως μόνο τα αντισώματα από ΧΛΛ περιπτώσεις με επιθετική νόσο εμφάνισαν σαφή αναγνώριση αντιγόνων αντιπυρηνικών αντισωμάτων. Επιπλέον, μόνο τα αντισώματα από κλινικά ήπιες ΧΛΛ περιπτώσεις αναγνώρισαν αντιγόνα στην κυτταρική επιφάνεια ζωντανών κυττάρων της Β λευχαιμικής κυτταρικής σειράς MEC1. Τέλος, παρατηρήσαμε ότι τα ΣΛΟΖ μονοκλωνικά αντισώματα αναγνωρίζουν πυρηνικά και κυτταροπλασματικά αντιγόνα, ένδειξη ότι οι ΣΛΟΖ Β κυτταρικοί κλώνοι πιθανόν προκύπτουν από την εξαλλαγή πολυαντιδραστικών Β λεμφοκυττάρων στην οριακή ζώνη του σπλήνα. Εντυπωσιακό είναι το γεγονός πως τα ΣΛΟΖ μονοκλωνικά αντισώματα έδειξαν σε γενικές γραμμές ισχυρότερη σύνδεση με δίκλωνο DNA και τάση για υψηλότερη αντιδραστικότητα έναντι αντιγόνων αντιπυρηνικών αντισωμάτων, συγκριτικά με τα ΧΛΛ μονοκλωνικά αντισώματα. Αυτή η παρατήρηση υποδηλώνει διαφορές σχετικά με το ιστορικό έκθεσης σε αντιγόνα και/ή τις επακόλουθες

ανοσοαποκρίσεις μεταξύ των καρκινικών Β λεμφοκυττάρων και/ή των προ-καρκινικών Β λεμφοκυττάρων μεταξύ των 2 διαφορετικών τύπων B-NHLs.

Abstract

Molecular studies of the clonotypic B cell Receptor (BCR) immunoglobulins (IGs) have contributed significantly to understanding the molecular basis of malignancies of mature B cells (B cell non-Hodgkin lymphomas, B-NHLs) and cemented the idea of antigen selection in disease ontogeny and evolution. However, despite recent progress, exactly which antigens are selecting and perhaps even continuously stimulating the malignant clones still remains to be elucidated.

The present study focused on 2 types of B-NHL, namely Chronic Lymphocytic Leukemia (CLL) and Splenic Marginal Zone Lymphoma (SMZL). Clonotypic BCR IGs were produced as recombinant IgM monoclonal Antibodies (mAbs) from 8 SMZL and 34 CLL patients. Recombinant mAbs were profiled for antigen reactivity using ELISA assays, flow cytometric analysis and Western blot analysis.

First, we noted that cases belonging to the same CLL stereotyped subset [subsets of cases expressing (quasi) identical BCRs] exhibited similar antigen reactivity profile. Interestingly, we also found that only mAbs from CLL cases with unmutated IGHV genes (U-CLL), generally associated with aggressive disease, exhibited a detectable anti-nuclear antibody (ANA) reactivity. Moreover, mAbs from indolent CLL cases with mutated IGHV genes but not aggressive U-CLL cases bound to the surfaces of viable leukemic MEC1 cells. In SMZL, our studies revealed recognition of nuclear and cytoplasmic antigens by SMZL mAbs expressing the most common IG gene rearrangements, thus suggesting that SMZL B cell clones may arise from polyreactive B cells in the splenic marginal zone. Interestingly, SMZL mAbs generally showed a stronger anti-dsDNA reactivity and a tendency for a higher ANA reactivity compared to CLL mAbs. This observation implies differences in the antigen exposure history and/or in the subsequent immune responses between the malignant B cells and/or the pre-malignant B cell progenitors of these two distinct entities.

Introduction

1. The Immune System: structure and function

Although the human body contains thousand species of microscopic organisms that constitute part of its normal flora, many microbes can become pathogenic and cause disease. The immune system, a host defense system consisting of many biological structures and processes within the body, is uniquely poised to recognize a variety of antigens i.e. molecules capable of inducing an immune response. Antigens can generally be classified as exogenous antigens and endogenous antigens or autoantigens.

The immune system is divided into an innate and an adaptive branch. The former provides the first line of defense against foreign molecules, does not require prior exposure to the microbes and responds in a generic way. However, it does not provide long-lasting immunity to the host, since there is little or no memory. Epithelial barriers and leukocytes (neutrophils, macrophages, NK cells, lymphocytes and mast cells) comprise the cellular components of innate immunity and a limited number of germline-encoded receptors are responsible for innate immune recognition.[1] On the other hand, the adaptive immune system is highly specific for a particular pathogen and can provide long-lasting immunity. The cellular components of adaptive immunity are B and T lymphocytes and the receptors that are specific for the pathogens are not germline-encoded but rather acquired during the lifetime of the organism.[2]

Lymphocytes are produced by hematopoietic pluripotent stem cells in the bone marrow, which give rise to Common Myeloid Progenitors (CMP) and Common Lymphoid Progenitors (CLP).[3] The former generate the myeloid lineages -including monocytes, neutrophils, eosinophils and basophils- whereas the latter generate the lymphoid lineage -including B lymphocytes, T lymphocytes and natural killer NK cells.[4] Hematopoietic stem cells are controlled by local tissue microenvironments in the bone marrow called “niches” while the differentiation of their progeny is regulated by a variety of transcription factors.[5]

T and B lymphocytes, the major cellular components of the adaptive immunity and NK cells, that mainly kill virally infected cells and cancer cells, reside in the organs of the lymphoid system.[6] Primary or central lymphoid organs include the bone marrow and thymus, whereas secondary or peripheral lymphoid organs include lymph nodes, spleen, mucosa associated lymphoid tissue (MALT), tonsils and Peyer's patches.[7]

1.1 Innate Immunity

The term innate immunity refers to the first line of defense that gets activated immediately or almost immediately after the organism senses the presence of antigens. Innate immunity acts in a generic way against any kind of infection and includes a number of physical barriers, defense mechanisms and general responses. Physical barriers such as the skin, the gastrointestinal tract, the respiratory tract and cilia are continuously working in order to block or slow the microbial invasion, whereas defense mechanisms such as secretions (mucous, bile, gastric acid, saliva, tears, and sweat) may trap and wash the microbes away.[8] In addition, general responses include: 1) inflammation, where immune cells are actively brought to the site of infection due to increased blood flow, 2) phagocytosis, where cells internalize microbial pathogens and 3) complement, which acts in several ways to mobilize defense mechanisms.[8]

Processes of the innate immunity are directed by a limited number of germline-encoded receptors, known as Pattern Recognition Receptors (PRRs), which are thought to have evolved prior to the emergence of adaptive immunity. These are proteins expressed by cells of innate immunity and are capable of recognizing two kinds of molecules: 1) Pathogen-Associated Molecular Patterns (PAMPs), which are structurally conserved motifs within a class of microbes; and 2) Damage-Associated Molecular Patterns (DAMPs), which are intracellular molecules that remain unrecognized by the immune system under normal conditions but are released under conditions of cellular stress or tissue injury.[9] A variety of different types of molecules from bacteria, fungi, parasites and viruses such as Bacterial Lipopolysaccharides (LPSs), endotoxins, bacterial flagellin, lipoteichoic acid, peptidoglycan, and nucleic acid variants [e.g. single-stranded RNA (ssRNA), double-stranded RNA (dsRNA) or unmethylated CpG motifs] serve as PAMPs. Their recognition by PRRs rapidly triggers a series of anti-microbial immune responses.[9]

1.1.1 Toll like receptors

There are several subgroups of PRRs which, according to their localization, can be designated as membrane-bound PRRs, which include mainly the Toll like receptors-TLRs, or cytoplasmic PRRs, which include mainly the NOD-like receptors -NLRs. TLRs were the first PRRs identified in mammals and are still the best characterized. In humans a total of 10 different TLRs have been identified.[10] TLRs are type I transmembrane proteins and can be divided into two subgroups according to their cellular localization and the respective PAMP ligands. Microbial components such as lipids, lipoproteins and proteins are identified by the first subgroup of TLRs namely TLR1, TLR2, TLR4, TLR5 and TLR6, which are expressed on the cell surface, in contrast with microbial nucleic acids which are identified by the second subgroup of TLRs (TLR3, TLR7, TLR8, TLR9 and TLR10) that reside in intracellular vesicles.[10] Besides initiating crucial inflammatory responses, TLRs also shape adaptive immunity, thus constituting an important link between the innate and the adaptive immune system.[11]

1.1.2 NOD-like receptors

NLRs are intracellular cytosolic protein-sensors: 23 NLR family members have been identified in humans.[12] NLRs are expressed in many cell types such as immune and epithelial cells and they recognize both PAMPs and DAMPs. Moreover, they organize signaling complexes such as NOD signalosomes, which trigger inflammatory responses and inflammasomes that activate inflammatory caspases and the highly pro-inflammatory cytokines of the interleukin-1 family .[12]

1.2 Adaptive immunity

Two main features distinguish adaptive from innate immune responses: ‘specificity’ and ‘memory’. Specificity refers to the ability of the immune system to recognize and respond to a particular antigen amongst a number of different antigens, whereas memory means that pathogens are "remembered"; thus upon a future exposure, a stronger and faster immune response is elicited.[8]

Adaptive immunity is classified into antibody-mediated or humoral and cell-mediated, mediated by B and T lymphocytes, respectively. Adaptive responses rely on clonal expansion of B and T lymphocytes which carry B cell Receptors (BcR) and T cell Receptors (TR), respectively, and can be divided into three phases:

1) 'Recognition phase' where the antigen gets recognized by Antigen Presenting Cells (APCs) which process and present the antigen to T helper lymphocytes.
2) 'Activation phase' where lymphocytes get activated, proliferate and differentiate mainly into effector cells and, to a lesser extent, into memory cells. This is known as clonal expansion.
3) 'Effector phase' where the foreign antigen is eliminated by either activated cytotoxic T cells (cell-mediated immunity) or antibodies released by activated B cells (humoral immunity).

1.2.1 Humoral immunity

During humoral responses, B cells recognize a foreign antigen that circulates in the lymph or blood, leading to their proliferation and generation of plasma cells, which secrete antibodies specific for that particular antigen.[13] Antibodies exert their function in three distinct ways:

1) Neutralization of pathogens and their toxins, by attaching to parts of their surface, thus blocking their ability to bind to receptors on host cells.[1,13]
2) Opsonization which means that antibodies latch onto the surface of pathogens and enhance their phagocytosis. [13]
3) Complement activation, in which antibodies attached onto the surface of foreign molecules stimulate complement system through a cascade of reactions. [1,13]

1.2.2 T cell-mediated immunity

Some pathogens, such as viruses, are not accessible to antibodies since they replicate inside the cells, thus cell-mediated responses are necessary for their elimination. Cell mediated responses are mainly tailored by T cells which are subdivided into cytotoxic T cells (Tc), that express the CD8 glycoprotein on their cell surface, and helper T cells

(Th) that express the CD4 glycoprotein.[2] T cells recognize short antigenic peptides bound to major histocompatibility complex (MHC) class I or II molecules.

Any cell (virus-infected, tumor, or transplanted) displaying MHC-I proteins bound to antigens on its surface triggers the activation and subsequent proliferation of cytotoxic Tc cells that cause cell destruction.[8] On the other hand, antigens coupled with MHC-II displayed on the cell membrane trigger the activation of Th cells, which stimulate B cells to produce antibodies that, in turn, bind to the antigens and recruit other types of cells (NK and macrophages) to destroy them.[8]

In conclusion, upon exposure to a foreign molecule, innate responses are the first that come to the foreground and adaptive immunity comes to complement and reinforce the total procedure.

2. Antigens

2.1 General

Infectious agents contain a variety of substances that are called immunogens or antigens, capable of eliciting an immune response. Nonetheless, there is a distinction between the terms "immunogen" and "antigen". The former term refers to any substance capable of inducing a humoral and/or cellular immune response, whereas the latter refers to any substance that can bind to the products of immune responses, including antibodies and TRs. Hence, an immunogen is necessarily an antigen, but an antigen may not necessarily be an immunogen. For simplicity, in the current postgraduate diploma thesis, the term antigen will be used when referring to both antigens and immunogens.

Lymphocytes recognize and bind to specific areas of macromolecules, known as epitopes or antigenic determinants. An epitope or antigenic determinant is any configuration on the surface of the antigen that can bind to its cognate antibody or TR. Due to the fact that almost any formation on the surface of a macromolecule can act as an epitope, any given structure (single molecule or supramolecular complex) usually contains a large number of epitopes. Antibodies bind to a specific epitope with different strength, which determines the affinity of the antibody for this epitope. Antibody avidity stands as a measure of the overall strength of an antibody-antigen complex.

Two different molecules may have one or more epitopes in common. As a result, products generated against an epitope of one molecule,can also recognize one or more epitopes on another molecule, a phenomenon called cross-reactivity.[14]

2.2 Properties of antigens

Several factors influence the immunogenicity of a given substance. In particular, an immunogen must be recognized as non-self by the organism and, generally, the greater its molecular size, the greater its immunogenicity. Moreover, molecules with more complex chemical structure or the ability to be degraded by antigen presenting cells are strong immunogens. Antigen dosage as well as the route of administration are important factors too, since above or below a certain dose the immune response is not optimal. Moreover, the subcutaneous route of administration increases the immunogenicity of a substance, compared to the intravenous or intragastric route.

Finally, the chemical nature of the antigens affects their immunogenicity. For example, proteins and polysaccharides have strong immunogenic properties in contrast to lipids and nucleic acids, which can elicit immune responses when associated with a carrier protein.

2.3 Types of antigens

Antigens can be classified according to their source into exogenous or endogenous. Exogenous antigens enter the body from the outside, whereas endogenous antigens are generated within normal cells as a result of normal cell metabolism, or due to viral/bacterial infection.

The BcR recognizes and binds to epitopes on the surface of soluble antigens. The humoral immune response begins either when the B cell recognizes an epitope and receives a co-stimulatory signal from an active Th cell or directly from microbial components without the involvement of T cells. Based on this distinction, antigens can be classified into two categories: Thymus Dependent (TD) and Thymus Independent (TI).[8]

Antigen-presenting cells (APCs) engulf and digest pathogens into smaller peptides containing epitopes, which are then presented, complexed with major histocompatibility complex (MHC) molecules, to T cells. These epitopes constitute

linear hydrophobic regions of proteins. Polysaccharides do not have such regions; thus, they can be recognized by B cells but not by T cells (TI antigens). On the contrary, proteins possess epitopes that can be recognized by both B and T cells (TD antigens).[8]

In TI responses, B cells can be activated by direct BcR crosslinking as well as through the TLRs. TI antigens can be divided into two subgroups: 1) TI-1 antigens such as bacterial lipopolysaccharides (LPS), that are completely independent of T cells and, at high concentrations, have the potential of polyclonal (non-specific) B cell activation and 2) TI-2 antigens such as ficoll, that consist of a highly repetitive surface structure, do not possess the ability of polyclonal B activation and require a small number of cytokines from Th cells.[8]

Autoantigens are normal proteins or DNA/RNA that under normal conditions are not recognized as foreign, but in cases of autoimmunity may elicit immune responses.

Finally, superantigens constitute a class of antigens that can cause polyclonal T cell activation followed by massive cytokine release. They bind first to MHC Class II expressed on APCs and then to the variable region of the TR β-chain. These antigens are generated by some viruses and bacteria and are capable of activating up to 20% of total T cells.[15]

3. Antibodies

3.1 Structure and function

Humoral immunity is based on antibodies (Abs), also known as immunoglobulins (IGs), which are glycoproteins produced by B cells that exert two main functions: they bind to the epitopes of pathogens that have triggered their production and activate the mechanisms responsible for their elimination.

Antibodies are Y-shaped molecules that consist of four polypeptide chains, two identical heavy chains and two identical light chains, connected by disulfide bonds.[16] (Figure 1)

Mammalian antibodies are classified into five main classes, known as isotypes, according to the type of heavy chain that they contain. An IG has α, γ, δ, ε or μ heavy chains that define its class: immunoglobulin A (IgA), immunoglobulin G (IgG),

immunoglobulin D (IgD), immunoglobulin E (IgE) or immunoglobulin M (IgM), respectively. IgG and IgA are further subdivided into IgG1, IgG2, IgG3, IgG4 and IgA1, IgA2 subclasses. In mammals there are two types of IG light chains, kappa (κ) and lambda (λ). Each IG carries light chains of only one type.

IgG, IgM and IgA are the most abundant IGs, whereas IgD and IgE comprise less than 1% of the total amount.[8] IgM is secreted as a pentamer and IgA is secreted as either a dimer or a trimer of the basic functional unit of antibodies, which is the IG monomer.[8]

In vitro an IG monomer can be cleaved by the enzyme papain into two Fab fragments and an Fc fragment. Fab (antigen-binding fragment) is the region of the antibody that has one constant and one variable domain of each heavy and light chain, whereas Fc (Fragment, crystallizable) is the region of the antibody consisting of two heavy chains that, depending on the class of the antibody, confer two or three constant domains.[17] Fc region was given this name because it was observed to crystallize readily.

The structure of an antibody is related to its function. Every chain and, therefore, every antibody molecule is subdivided into the Variable domain at the N-terminus and the Constant domain at the C-terminus, which define the specificity of the antibody (recognition of the antigen) and its biological function, respectively.[8] Each antibody chain has a tandem series of repeating homology units of about 70–110 amino acids in length, called IG domains. Each domain consist of two β-pleated sheet structures held together by interactions between conserved cysteines and other charged amino acids. [18] The constant region of α, γ and δ heavy chains includes three domains (CH1, CH2, CH3), whereas the constant region of μ and ε heavy chains includes four domains (CH1, CH2, CH3, CH4).[8] These multiple CH domains are numbered from the N-terminus to the C-terminus. On the contrary, the Variable domain of heavy and light chains has only one IG domain (VH or VL) and is identical only for those antibodies that derive from a single B cell or B cell clone.

The amino acid sequences that span at the N-terminus of heavy and light chains show great variability between different antibodies.[8] However, this variability is not uniformly distributed, since there are regions that exhibit great variability and other regions that are more conserved. The former are called hypervariable regions or Complementarity Determining Regions (CDRs) and correspond to three loops at the end of each beta sheet, both in the heavy and the light chain.[8,19] During protein folding,

they come to close proximity and generate the antigen-binding site (paratope).[8] There are three CDRs: CDR1, CDR2 and CDR3. On the other hand, the more conserved regions between CDRs are designated as Framework Regions (FR) because they provide the structural framework of the domain.[8]

The amino acid sequences that span at the C-terminus of heavy and light chains do not display the same diversity and are called the Constant (C) domain. The C domain of the antibody determines its biological function (the antibody class effect) with different isotypes of antibodies having different class effects. These include antibody-dependent cellular cytotoxicity, activation of the complement system and removal of antigens from macrophages.

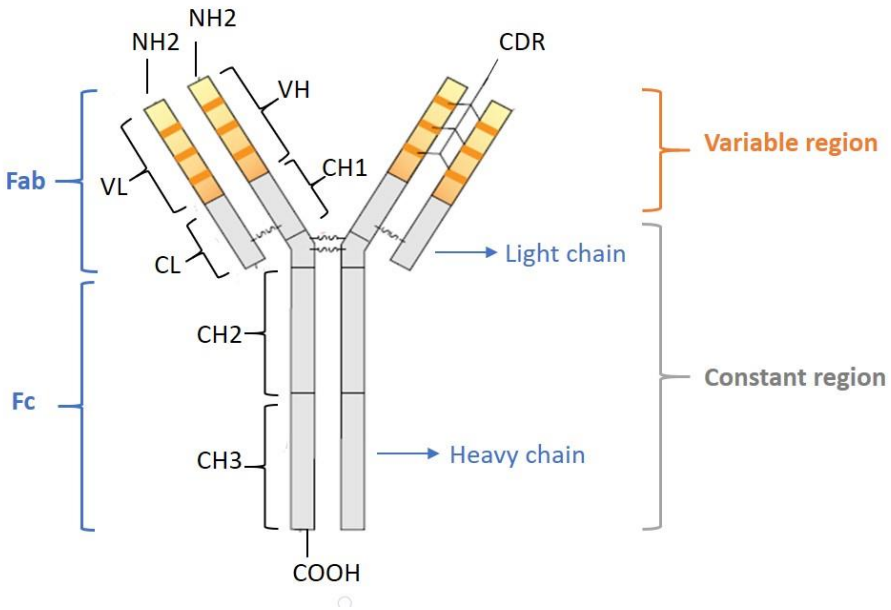


Figure 1: *The basic unit of an immunoglobulin, composed of two identical heavy chains and two identical light chains. Both the heavy chain and the light chain include a variable and a constant region.*

3.2 Immunoglobulin Genes

The heavy chain and light chain polypeptides are produced by separate IG genes: V (Variable), J (Junctional), and C (Constant). The IG heavy chain gene locus (IGH) includes also the D (Diversity) genes between the V and J genes. One gene from each group is randomly selected for the synthesis of the variable domains of IGs.

IG gene loci also include a number of regulatory genes, such as leader sequences (L) before each V gene, which direct the newly-formed heavy and light chains into the endoplasmic reticulum.

In humans, the IGH locus is located on chromosome 14q32.33 and spans 1250 kilobases (kb). Per haploid genome, 170 to 176 heavy immunoglobulin genes exist, of which 77 to 84 are functional.[20] In more detail, there are 44 functional IGHV, 23 functional IGHD, 6 functional IGHJ and 11 functional IGHC genes. IGHV and IGHJ groups are also divided into 7 subgroups: IGHV1-IGHV7 and IGHJ1- IGHJ7, respectively.[20] Some IGH genes cannot participate in the assembly of IG polypeptide chains, albeit they have an Open Reading Frame (ORF). These are called orphons, they are located outside the main locus and thirty-five of these genes have been identified.[20]

The IGHV genes in humans are classified into three phylogenetic clans (I, II, and III):

- clan I consists of IGHV1, IGHV5 and IGHV7 subgroup genes
- clan II consists of IGHV2, IGHV4 and IGHV6 subgroup genes, as well as pseudogenes IGHV(II) and
- clan III consists of IGHV3 subgroup genes and pseudogenes IGHV(III)

The Human Immunoglobulin Kappa (IGK) gene locus is located on chromosome 2p11.2 and spans 1820 kb. It includes 76 IGKV genes (of which 31-35 are functional), 5 IGKJ genes and a unique IGKC gene.[21] The IGKV genes belong to 7 subgroups (IGKV1-IGKV7) and are organized in two clusters that are separated by 800 kb. Moreover, 28 IGKV orphons have been identified.

Finally, the Human Immunoglobulin lambda (IGL) gene locus is located on chromosome 22q11.2 and spans 1050 kb. It includes 73-74 IGLV genes, 7-11 IGLJ and 7-11 IGLC genes, depending on the haplotypes, and 14 pseudogenes. The potential IGL repertoire consists of 29-32 functional IGLV genes belonging to 10 subgroups, 4-5 IGLJ and 4-5 IGLC functional genes.[22]

3.3 Mechanisms of immunoglobulin diversity

IGs are able to recognize a great number of antigens in order to protect the organism from various threats. Virtually any substance can elicit an antibody response. IG diversity rests on the combined effect of specific mechanisms that function during B cell development:

1) Combinatorial diversification: The large number of germline V, D and J genes contribute to the large diversity, since any of the V genes can be joined to any of the D (for IG heavy chains only) and J genes. Moreover, the assembly of a heavy chain with a light chain (k or λ) increases the combinatorial diversity. (Figure 2)
2) Junctional diversification: During V(D)J recombination, diversity is enhanced at the junctions of V, (D) and J genes because nucleotides may be trimmed from the ends of the recombined genes and random (without template) nucleotides may be added. (Figure 2)
3) Somatic hypermutation (SHM): Following activation by antigen, B cells begin to proliferate rapidly. In these rapidly dividing cells, the genes encoding the variable domains of the heavy and light chains undergo a remarkably high rate of somatic mutation as a result of a process called somatic hypermutation (SHM) which increases antibody diversity and produces antibodies with higher specificity. (Figure 2)
4) Class switch recombination (CSR): In the IG locus, IGHM and IGHD are the first two Constant domain genes expressed, thus naïve mature B cells produce both IgM and IgD. When a B cell gets activated by a specific antigen, it undergoes class switch recombination to produce antibodies of different isotype (IgG, IgA or IgE). Although the constant domain of the IG heavy chain changes, the variable domain remains the same, thus the specificity for the antigen is not influenced. This means that the antibody is still capable of recognizing the specific antigen, but it acquires different biological functions.

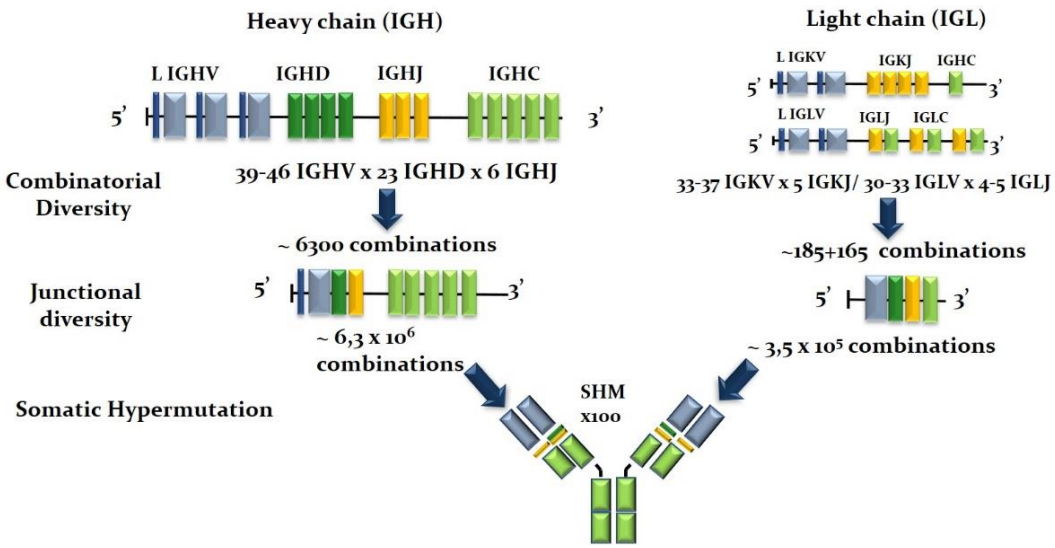


Figure 2: *Mechanisms underlying immunoglobulin diversity: 1) Random recombination of immunoglobulin heavy and light chain genes, 2) random addition of nucleotides to the junctions of recombining genes by TdT, 3) random assortment of heavy chains with light chains and 4) affinity maturation through somatic hypermutation upon antigen exposure.*

3.3.1 Combinatorial diversification and V(D)J recombination

An antibody consists of four polypeptide chains, two identical heavy and two identical light chains. The gene sequences for variable domain of each chain are formed from the random assortment of a gene from the V, D and J genes. The various available genes in each group result in an extremely high number of combinations.

A process called “somatic recombination” or “V(D)J recombination” takes place during the early phase of B cell differentiation in the bone marrow, in order to generate the variable domain of IGs. This procedure is directed by flanking DNA sequences called Recombinant Signal Sequences (RSS), which consist of a conserved heptamer (5´-CACAGTG-3´) and a conserved nonamer (5´-ACAAAAACC-3´) sequence, separated by a spacer region.[23,24] The latter is either 12 base pairs or 23 base pairs long and plays a crucial role in preventing the recombination of two different genes coding for the same region (e.g. V-V recombination). In more detail, genes belonging to the same gene group (e.g. V group) have the same type of RSS (same length of spacer region), which are located at the 3 'end of the V genes, at the 5' end of the J genes and at the 5 'and 3' ends of the D genes. Recombination is possible only between genes having RSS with different length of spacer regions, meaning that a 12 base pair spacer RSS can be

recombined only with a 23 base pair spacer RSS and vice versa. This rule is known as "12/23 rule" or "one turn/two turns" because 12 and 23 base pairs correspond to one or two turns of the double helix of DNA, respectively, orienting the heptameric and nonameric sequences on the same side of the molecule.[23,24] (Figure 3)

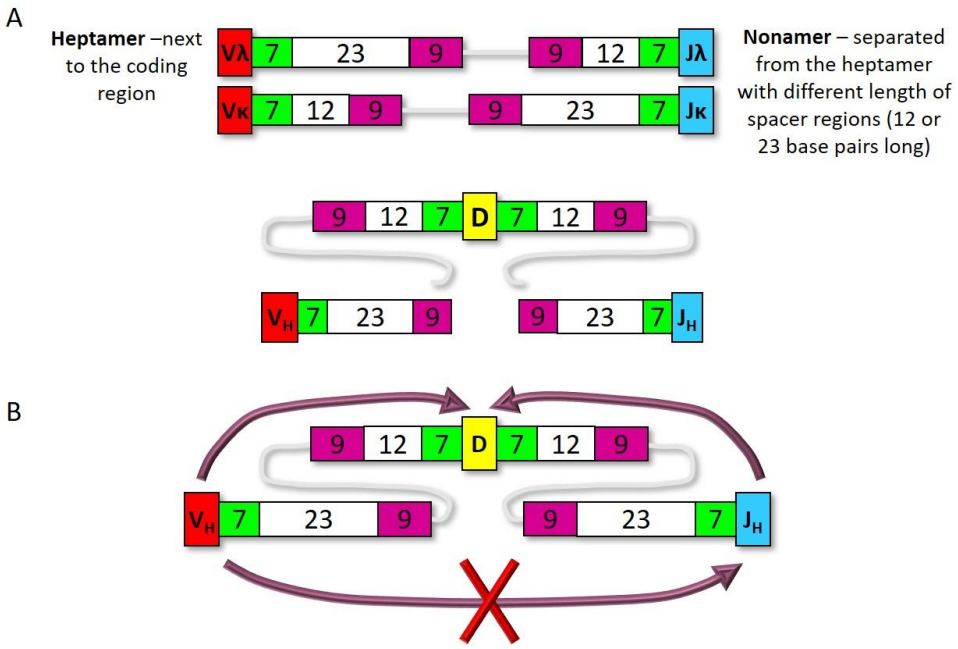


Figure 3: *A) V(D)J recombination is directed by flanking DNA sequences called Recombinant Signal Sequences (RSS), which consist of a conserved heptamer and a conserved nonamer sequence, separated by a spacer region of 12 or 23 base pairs long. Genes belonging to the same gene group (e.g. V group) have the same type of RSS (same length of spacer region). B) The 12/23 rule: Recombination is possible only between genes having RSS with spacer regions of different length.*

At the biochemical level, V(D)J recombination is divided into two stages:

1) cleavage of the DNA and
2) processing and joining of the cleaved DNA ends

The first stage requires the RAG1 and RAG2 proteins, which are enzymes encoded by the recombination-activating genes (RAGs). RAG1 and RAG2 alone can bind DNA, but they bind an RSS more tightly and specifically when they form a complex together.[25] First, a nick is made at the 5' end of the RSS heptamer, resulting in the formation of a free 5'-phosphate (PO_4) group at the RSS end and a free 3'-hydroxyl (OH) group at the coding end. Then, the reactive 3 -hydroxyl group attacks the opposite

strand and binds to the 5'- PO_4 group (which is located between the RSS and the gene segment of the opposite strand), giving rise to a 5'-phosphorylated double-stranded break at the RSS and a covalently closed hairpin at the coding end.[25]

The second stage requires a number of enzymes, namely DNA-PK, Ku, XRCC4 and DNA ligase IV, in order to join the cleaved DNA ends through the Non-Homologous End Joining (NHEJ) pathway. The signal ends are linked to each other by DNA ligase IV. The coding ends are processed further prior to their ligation.[24] In particular, they are opened by the activity of Artemis that, in turn, gets activated by DNA-PK [a complex consisting of the catalytic subunit (DNAPKCS) and the heterodimer Ku protein (Ku70 and Ku80)].[24] The coding ends are further processed by addition and deletion of nucleotides and finally, are ligated together by DNA ligase IV. DNA ligase IV is one of the three genetically distinct ligases in mammalian cells and XRCC4 is crucial for the NHEJ ligation step, since it enhances the joining activity of DNA ligase IV. [26,27]

Following V(D)J recombination, transcription of the recombinant genes, slicing of the precursor mRNA (pre-mRNA) and addition of 3' poly-A tail take place, in order to give the final product.

Control of V(D)J recombination

V(D)J recombination is tightly controlled by several mechanisms:

- At the transcriptional level, the expression of RAG proteins is limited to precursor B and T cells. Furthermore, there are restrictions regarding the lymphoid cell type (B,T) and the developmental stage, meaning that although the same recombination mechanism exists, IG genes become fully rearranged only in B cells, and TR genes only in T cells.[23,24] Moreover, the recombination machinery is limited to specific sites of DNA (Recombination Signal Sequences) reducing the available boundless combinations (regulation of site accessibility).
- Recombination occurs in a hierarchical pattern. In B cells, the IG Heavy chain genes are rearranged before the light chain genes. In addition, kappa light chain gene rearrangement precedes that of lambda light chain.[28]
- Successful rearrangement of heavy chain genes on one chromosome prevents rearrangement on the other allele and this limitation is known as allelic exclusion.

However, if rearrangement on the first allele is not successful, rearrangement on the second chromosome occurs. Unsuccessful rearrangement on both chromosomes leads to cell death.[29]

3.3.2 Junctional diversity

In both heavy and light chains, the two hypervariable regions (CDR1 and CDR2) are encoded exclusively by the V genes. In contrast, the CDR3 region includes the V-D-J and V-J genes for the heavy and light chains, respectively. The imprecise joining of V, (D) and J gene segments during V(D)J recombination creates even greater diversity in the CDR3 region, which is caused by the deletion of nucleotides in the joint region and the addition of the so-called P and N nucleotides (Figure 3).

The two mechanisms mentioned above operate during V(D)J rearrangement in developing B cells occurring in the bone marrow. However, the final maturation of B cells takes place in the Germinal Centers (GC) of the secondary lymphoid organs, following antigen recognition, and includes two additional mechanisms of diversity generation: 1) Somatic hypermutation (SHM) and 2) Class Switch Recombination (CSR)

3.3.3 Somatic hypermutation

When B cells encounter antigen, a process called Somatic Hypermutation (SHM) begins, leading to the introduction of point mutations into the variable regions of the rearranged heavy and light chain IG genes. Some of the altered IGs generated by SHM possess enhanced ability to recognize and bind the specific antigen (higher affinity) and, thus, are preferentially stimulated by the antigen to survive and proliferate. B cells with very high affinity antibodies derive from a series of mutation and selection cycles, a procedure known as affinity maturation.[30]

SHM is divided in two phases: (i) the Activation-induced cytidine deaminase (AID) deaminates cytidine residues to deoxyuridine residues, thus generating a DNA lesion and (ii) error-prone repair mechanisms process the lesions, resulting in introduction of point mutations.[31,32] In more detail, the cytosine:guanine pair produced by AID is mutated to an uracil:guanine mismatch. Subsequently, high-fidelity “base excision repair enzymes” are recruited to repair these mutations, since uracil residues are not

normal components of DNA. Uracil-DNA glycosylase removes the uracil residues and error-prone DNA polymerases are then recruited to fill the gap, resulting in new mutations.[31,33]

While, the frequency of basal mutation level in the genome is ~10^{-9} per base pair per generation, the mutation frequency in the rearranged V region is approximately 10^{-5}-10^{-3}/base pair/generation. These mutations preferentially occur within specific oligonudeotide motifs such as RGYW, WRCY and GNW (R: purine, Y: pyrimidine, W: A or T), known as "hotspots".[34–36] Moreover, there are more transitions (e.g., C to T, G to A) than transversions (e.g., C to A or G; G to C or T) and ~60% of the mutations are in A:T bases.[36]

3.3.4 Class switch recombination

Mature B cells that have encountered antigen, can change their IG heavy constant domain, leading to different effector function. Thus, different daughter cells from the same activated parent B cell are capable of producing antibodies with identical variable domain and different IG heavy constant domain, meaning that they retain their antigen specificity but can interact with different effector molecules. The replacement of the constant (IGHC) gene to be expressed from IGHM to IGHG or IGHE or IGHA, switching antibody production from IgM to a different class such as IgG, IgE or IgA, is called Class Switch Recombination (CSR) and the different categories of antibodies generated by the altered heavy constant segment are known as isotype.

In humans, the IGHC gene cluster is located on chromosome 14. It includes nine functional genes as well as two pseudogenes. The IGHM and IGHD genes are transcribed from the same transcriptional unit, thus naive mature B cells display both IgM and IgD on their surface.[37] The mRNA produced, through alternative transcriptional termination and splicing, gives rise to IGs of isotype M and D that have exactly the same recombined VDJ gene complex. However, after B cell activation by antigen, expression of isotype D IG stops and cells either secrete IgM or switch to another isotype through CSR.[37] CSR is an irreversible process achieved by recombination between specific "switch regions" (S regions) that are located ~2 kb

upstream from gene segments that encode the IGHC (except for Cd) and consist of tandem repeats that differ in sequence and length.[38]

A number of cytokines and B cell activators mediate CSR by directing the transcription of non-rearranged IGHC genes prior to CSR, generating the so-called germline (GL) transcripts.[39,40] Interaction of CD40 BcR with the CD40L ligand of a T cell in the presence of cytokines influences the transcription which begins from an I (intervening) exon upstream of the S region. Transcripts are spliced to the first exon of the corresponding IGHC gene and the intervening sequences are excised as circular DNA.[41] Cytokines activate the promoters present at the 5' end of S regions, leading to the production of these non-coding transcripts that direct CSR procedure on DNA level.

Activation-induced deaminase (AID) plays a crucial role, probably by deaminating the dC residues within the S regions.[42] Subsequently, double strand breaks are introduced at the two selected S-regions and a number of DNA damage response/repair pathways is recruited, including ataxia-telangiectasia mutated (ATM)/ataxia-telangiectasia and Rad3-related (ATR)-dependent signaling, base excision repair (BER), mismatch repair (MMR), and nonhomologous end joining (NHEJ), resulting in the recombination of regions involved.[43]

4. B lymphocytes

4.1 Maturation of B lymphocytes in the bone marrow

B cell development encompasses a continuum of stages that begin in primary lymphoid organs, with subsequent functional maturation in secondary lymphoid organs. Hematopoietic stem cells (HSC) give rise to B cells through a process that occurs in the fetal liver before birth and in the bone marrow afterwards.[44] Functional B cells arise from various differentiation stages, each of which is determined by the rearrangements of IG genes and the expression of a number of molecules as well as growth factor receptors on their surface.[44,45]

4.2 Rearrangements of immunoglobulin genes

In CLP, rearrangement of IG genes has not yet occurred and RAG1 and RAG2 genes are expressed in low levels. The IG gene rearrangements in B-lineage cells follow a strict developmental order:

- ***Early pro B cell***: activation of IGH locus leads to recombination of IGHD and IGHJ genes on both alleles of heavy chain.
- ***Late pro-B cell***: the rearranged DJ complex recombines with an IGHV gene, to give a complete VDJ exon that encodes the variable domain of an IG heavy chain. In contrast to the IGHD-to-IGHJ recombination, IGHV to IGHD-IGHJ rearrangement takes place only on one allele. This is achieved by silencing of one of two homologous antigen receptor alleles, a process known as allelic exclusion. If the IGHV-IGHD-IGHJ rearrangement on the first allele is unproductive, rearrangement of the second IGH allele takes place.[29]

  The transition from IGHD-IGHJ to IGHV-to IGHD-IGHJ rearrangement, requires the presence of IL-7, Pax5 and YY1 transcription factors as well as high levels of RAG1/RAG2 proteins.[46] Further maturation of pro-B cells depends on the successful recombination and expression of the heavy chain IG genes. Cells that fail to give a functional IGHV-IGHD-IGHJ exon undergo apoptosis and phagocytosis by macrophages present in the bone marrow.[47] It is estimated that only 1/3 of all pro-B lymphocytes undergo IGHV-IGHD-IGHJ recombination leading to a functional product which allows transition to the late pre-B cell stage.[48]
- ***Early pre-B cell***: at this stage, the heavy chain is detected in the cytoplasm, whereas light chain genes have not yet been rearranged. On the cell surface, pre-B receptor (pre-BcR) is expressed, which consists of the transmembrane form of the μ heavy chain, a surrogate light chain (SLC) and the co-receptors Igα and Igβ.[49,50] The surrogate light chain is encoded by the VpreB and λ5 genes which are not subject to any rearrangement and are homologous to the variable and constant domain of the λ light chain, respectively.[51]
  - The newly synthesized heavy chains are tested for their ability to bind to the "surrogate" light chain (SLC).[52] Even a functional IGHV-IGHD-IGHJ rearrangement may result in a heavy chain unable to pair properly with the

surrogate light chain and it is estimated that only half of the produced heavy chains can successfully bind to the SLC.[53,54]

- Expression of a functional pre-BCR is required for further differentiation of B cells.[54] Pre-BcR transmits multiple signals to the developing B cell, resulting in protection from apoptosis, proliferation and transient downregulation of the expression of various genes, including RAG1/RAG2 and TdT.[55]

- ***Late pre-B cell***: during this step, expression of the pre-BcR and cell proliferation cease, whereas genes involved in the recombination that had been temporarily suppressed get activated for the rearrangement of light chain genes.
  - Allelic exclusion in IG light chains is achieved through limited accessibility of the recombination mechanism to the genetic loci of light chains. Rearrangements of light chain IG genes occur in a hierarchical manner and IGK gene locus is the first to rearrange.[29] Each newly synthesized light chain is tested for its ability to assemble with the existing heavy chain. If this proves impossible on both IGK alleles, recombination in IGL genes occurs.
- ***Immature B cell***: when a productive rearrangement for the light chain occurs, the complete BcR consisting of a dimer of IG heavy μ and light chain pairs, associated non-covalently with a heterodimer of Iga and Igb co-receptors, is expressed on the surface of the B cell. The cell at this stage is capable of recognizing antigens and may undergo both positive selection to give long-lived peripheral B cells and negative selection to eliminate potentially autoreactive cells.[56]

4.3 Central tolerance

Ligation of the BcR of the newly formed immature B cells present in the bone marrow by self-antigens promotes signaling that triggers regulatory processes to reduce self-reactivity. These processes are collectively known as central tolerance:

1) ***Apoptosis (clonal deletion)***: immature B cells that react with a self-antigen with high avidity undergo apoptosis.[57]
2) ***Anergy***: immature B cells that react with a self-antigen with low avidity enter into a state of non-reactivity to subsequent stimulation. The so called anergic B

cells lose the expression of IgM on their surface (although they retain IgM in the cytoplasm) and express high levels of IgD. Although anergic B cells are allowed to migrate into the periphery, they cannot enter follicles and have a reduced life-span.[57]

3) ***Receptor editing***: B cells that interact strongly with self-antigens, have the chance to escape apoptosis and anergy, through receptor editing. Binding of the BcR to its corresponding ligand can result in reactivation of RAG1 and RAG2 proteins, thus further recombination can occur. Receptor editing occurs in light chains and is a very efficient process that can, theoretically, provide a new non autoreactive IG to any autoreactive immature B cell. [58,59]

5. Maturation of B lymphocytes in the periphery

B cells with no or low affinity for autoantigens, exit the bone marrow and move towards the spleen in order to go through stricter selection. During this intermediate stage in their development, B cells are called transitional and, besides increasing the expression of IgM on their surface, they also start producing surface IgD ($mIgM^{+++}mIgD^{+/-}$). [60] These cells will be selected to become either Marginal Zone B cells (MZ) or Follicular B cells (FO).

5.1 Follicular B cells

The majority of mature B cells are recirculating, found mostly in aggregates of B cells (called follicles) in secondary lymphoid organs and they are defined as follicular B cells. In the spleen, follicles are located next to T cell zones, thus upon an initial response, helper T cells and activated B cells can easily migrate toward one another and interact.[61] Therefore, follicular B cells are mainly involved in T cell-dependent responses, although there is a population that can respond in a T cell-independent manner to blood-borne pathogens.[62,63]

Formation of Germinal Center:

Most mature follicular B cells express high levels of IgD and low levels of IgM and are known as follicular type I B cells. When these cells encounter antigen, they receive

signals from T helper cells that enhance their proliferation and allow extensive interactions between T, B and dendritic cells.

In more detail, a number of stimulated cells migrate to the center of the follicles where they set up the formation of a germinal center, an organized structure consisting of a dark and a light zone. (Figure 4) At the initial state, B cells are present in the dark zone where they reduce the expression of surface IG and proliferate rapidly. During this step, B cells are called centroblasts and undergo extensive SHM and CSR.[64] Subsequently, B cells move to the light zone, where selection occurs regarding their affinity to antigen. Throughout this selection process, interactions with follicular dendritic cells (FDCs) and follicular T helper cells are crucial. B cells with the lowest affinity IG receptors undergo apoptosis, whereas B cells with the highest affinity IG receptors receive survival signals from T cells and differentiate into antibody-secreting cells (plasma cells) or memory B cells.[64]

Plasma cells are terminally differentiated cells that survive in the bone marrow and mucosal sites for a few weeks, limiting the time for the immune response. In contrast, memory B cells are long-lived resting cells, that do not produce antibodies. Nevertheless, they have undergone isotype switching and have a receptor with high affinity for their cognate antigen. Memory B cells respond quickly after any subsequent encounter with the same antigen.[65]

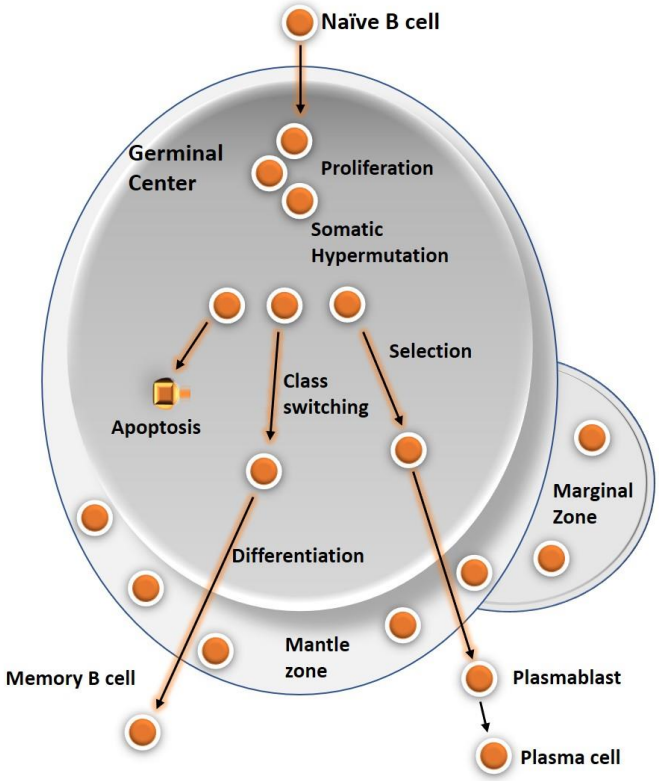


***Figure 4**: The main steps in mature B cell differentiation in the Germinal Center (GC). After antigen encounter, mature B cells are driven into a GC reaction, where GC B cells proliferate rapidly in the dark zone and activate the process of somatic hypermutation. In the light zone mutated B cells are selected for affinity-increasing IG V gene mutations, although most GC cells will acquire mutations and undergo apoptosis. Moreover, many GC cells undergo class switch recombination in the light zone. A clear marginal zone is present around B follicles in the spleen*

5.2 Splenic Marginal Zone B cells

The spleen, a highly vascularized organ, is anatomically divided into the red and the white pulp, which exert different functions. Red pulp consists mainly of vascular sinusoids in which blood circulates and macrophages present in this area remove microbes, especially non-opsonized particles and damaged cells. On the other hand, the white pulp consists of cells that mediate responses to blood-borne pathogens. In the white pulp, the central arterioles are surrounded by the periarteriolar lymphocyte sheaths (PALS), a T-lymphocyte-rich structure, where the outer PALS contain B cells that after antigenic stimulation become plasma cells. In addition. adjacent follicles contain B cells. Around each PALS and follicles is a corona containing B cells called the marginal zone. The marginal zone includes a distinct population of noncirculating mature B cells, the Marginal Zone B cells (MZ cells), which express high levels of surface IgM, CD21, CD1, CD9 with low to negligible levels of surface IgD. MZ cells are also present in tonsils and Peyer's plaques.[66]

MZ cells are mainly involved in immune responses to T-independent type 2 (TI-2) multivalent antigens with highly repetitive structures, such as bacterial polysaccharides.[67] After activation, they differentiate into short-lived plasma cells that produce mostly IgM. In humans, the marginal zone is immature before the age of two years, hence infants are unable to elicit effective immune responses against encapsulated bacteria.[68] Moreover, splenectomy in humans results in increased susceptibility to such microbes.[69] Although MZ B cells constitute only 5% of the splenic B cells, their role in early immune responses is of utmost importance because they represent a linkage between the innate immune responses, that are triggered within a few hours after antigen encounter, and the adaptive immune responses, which require almost two weeks.[66]

MZ B cells can be involved in T-dependent responses, too, since they express high levels of MHC class II antigens and B7 proteins, which are essential T cell co-stimulatory molecules.[70] In addition, MZ B cells may be subjected to CSR and SHM.[66]

5.3 B1 cells

B1 cells are characterized as a self-renewing population that express high levels of surface IgM and CD5, with low to negligible levels of surface IgD. They develop early during ontogeny in the fetus of rodents and are found in abundance in the peritoneum and mucosal sites. Some B1 cells that do not express CD5 on their surface exist and are known as B1-b cells.[71]

B1 cells spontaneously secrete antibodies in the absence of external antigen stimulation: such antibodies are preferentially of the IgM isotype and termed natural antibodies. Natural antibodies recognize self, altered self and foreign antigens consisting of phospholipids (e.g. phosphatidylcholine), carbohydrate sequences (e.g. gangliosides), single-stranded DNA and peptides (e.g. amyloid beta peptide) or surface glycoproteins (e.g. CD90).[72] They are considered to be the first line of defense against microbes and they rarely undergo SHM or CSR.[71]

B1 cells were originally studied in mice. In humans, their existence is controversial, although in a recent study a population of B cells with a CD20+ CD27+ CD43+ CD70- phenotype was identified as human B1–cells.[73,74]

6. B Cell Receptor Signaling

The B cell receptor (BcR) is a transmembrane protein complex on the surface of B cells, composed of membrane IG molecules and non-covalently associated Iga/Igb (CD79a/CD79b) heterodimers. Iga and Igb are members of the IG superfamily and contain a single extracellular IG-like domain, a single transmembrane region and a cytoplasmic tail containing an ITAM (Immunoreceptor Tyrosine-based Activation Motif) sequence responsible for signal transduction after BcR ligation.[75]

Membrane IG (mIG) molecules bind antigen resulting in receptor aggregation, which is followed by the phosphorylation of tyrosine residues in the ITAMs of Iga and Igb by kinases belonging to the Src family such as Fyn, Blk, Hck, Fgr and Lyn.[76] This phosphorylation is responsible for the recruitment of the Syk protein kinase which becomes activated due to phosphorylation by Src family protein kinases or autophosphorylation.[77] Syk activation starts the formation of the BcR signalosome, initiating a cascade of reactions, from which a number of relatively independent signaling systems diverge:

- RAS pathway induced AP-1 signaling
- PKC-beta-induced NF-kappaB signaling
- PI3K-Akt signaling
- Calcium release initiated NFAT activity and
- AVA-RAC mediated cytoskeletal rearrangement

The activation of two molecules, phosphatidylinositol 3-kinase (PI3K) and phospholipase Cγ2 (PLCγ2) is very crucial for the fate of B cells. PI3K leads to activation of the PI3K /Akt pathway, which promotes the survival of B cells by both inducing the expression of anti-apoptotic proteins as well as inactivating pro-apoptotic proteins.[78] On the other hand, activation of PLCγ2 triggers the release of intracellular Ca^{2+} with subsequent activation of protein kinase C (PKC). The cascade continues with the activation of Mitogen-Activated Protein Kinases (MAPKs), such as ERK (extracellular signal-regulated kinase), JNK (c-JUN NH2-terminal kinase) and p38 and the activation of transcription factors such as NF-κB (nuclear factor-κB) and NFAT (nuclear factor of activated T cells).[79] NF-κB is known to support cell-survival, whereas the aforementioned MAP kinases have been related to both proliferation and apoptosis of B cells.

7. Chronic Lymphocytic Leukemia

7.1 Introduction

Chronic lymphocytic leukemia (CLL) is a chronic B cell malignancy characterized by clonal expansion of $CD5^{+}CD23^{+}$ B cells in blood, bone marrow and second lymphoid tissues. It predominantly affects elderly individuals, with a median age at diagnosis of 72 years.[80] CLL is the most common adult leukemia in Western World with an incidence of 4.2/100 000/year. Of note, family studies have shown higher prevalence of the disease among relatives of patients with sporadic CLL, suggesting genetic predisposition.[81] Males and whites are more frequently affected than females and other races.[82–84] Moreover, incidence rates are lower in Asia and South America, but an explanation for this geographical variance does not exist.[85]

The diagnosis of CLL is established by the following criteria[86]

1) presence of a monoclonal B cell count of above 5000/μl in the peripheral blood, for a duration of at least 3 months.
2) The Immunophenotype of B CLL cells as assessed by flow cytometry should present the following features:
    a. co-expression of CD19, CD23, and CD5
    b. restriction to expression of either κ or λ immunoglobulin light chains, low levels of surface immunoglobulin, CD20 and CD79b.

Two staging systems for CLL are used to correlate clinical findings with survival times: the Rai and the Binet system.[87,88] Besides the clinical features, many molecular/biologic factors such as ZAP-70 and CD38 protein expression, IGHV gene SHM status, cytogenetic abnormalities and deletions/mutations of TP53 correlate with an inferior outcome.[85] Further studies are required to define new biomarkers which could be used for better stratifying patients towards personalized disease management.

7.2 Genomic aberrations

Chromosomal abnormalities can be detected in up to 80% of CLL patients and molecular cytogenetic analysis has revealed a number of lesions with important clinicobiological implications.[89,90] Deletions of chromosomes 11q, 13q, 17p, and trisomy 12 have a known prognostic value. In particular, deletion of 11q23 is associated with clinically progressive disease, whereas deletion of 13q14 has been associated with favorable prognosis.[90] Moreover, 17p13 deletion is one of the most frequently acquired aberrations, triggered after treatment and confers high risk for a more aggressive clinical course, maybe due to loss of the *TP53* tumor suppressor gene.[90] Finally, trisomy 12 is the third most frequent chromosomal abnormality and considered to be an intermediate risk marker.[90]

Ongoing advances in sequencing technology have revealed several key driver alterations in CLL cells. Mutations in *TP53* and *ATM* genes are the most common, whereas mutations in *MYD88, NOTCH1, XPO1, SF3B1* and *BIRC3* genes have also been observed.[91] These alterations affect a number of cellular processes: DNA damage and cell cycle control (*TP53, ATM, BIRC3*), mRNA processing (*XPO1, SF3B1*), NOTCH signaling (*NOTCH1*), inflammatory pathways (*MYD88*), as well as chromatin modification (*CHD2*).[92]

7.3 Immunogenetics in CLL

Immunogenetic analysis of the clonotypic IGHV-IGHD-IGHJ gene rearrangements can subdivide patients into two categories: mutated-CLL (M-CLL) and unmutated-CLL (U-CLL). This distinction is based on the presence or absence of SHM within the rearranged IGHV genes of the clonotypic BcR. Particularly, patients expressing clonally rearranged IGHV genes with >2% deviation from the germline sequence are assigned to the M-CLL category, whereas those with <2% belong to U-CLL.[93] The SHM status of IGHV genes represents a robust predictive biomarker, since U-CLL patients experience a significantly more aggressive disease than M-CLL.[94,95] (Figure 5)

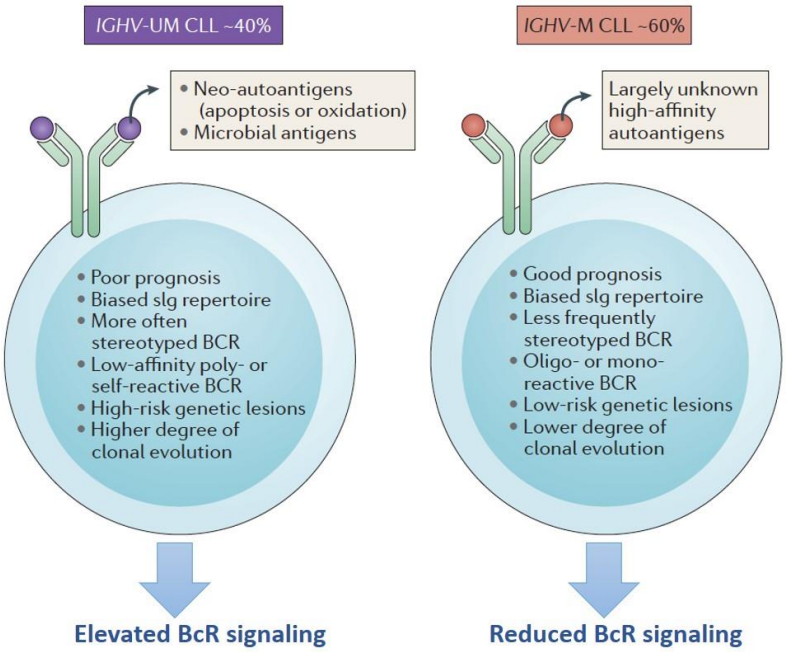


***Figure 5**: Distingushing features of IG-mutated vs IG-unmutated CLL. (Fabbri and Dalla-Favera. Nat Rev Cancer 2016)*

The presence of SHM in CLL mAbs constitutes a feature of cells that have passed through the germinal center of secondary lymphoid tissues, following antigen recognition.[96] However, absence of SHM does not necessarily mean absence of prior antigenic stimulation, since it may reflect the type of antigenic stimulation (TI antigens) or the stage at which the transition from normal to neoplastic state occurred.

A restriction in IGHV gene repertoire has been observed in CLL cells that differs from the normal repertoire of B cells. In particular, certain genes such as IGHV1-69, IGHV4-34, IGHV3-7 and IGHV3-23 are very common in CLL. (Figure 6) A preference regarding the mutational status has also been noticed:[97] for example, IGHV4-34 is more frequent in M-CLL cases, whereas IGHV1-69 in U-CLL.[97,98] This biased gene usage strongly supports the notion of antigen selection in disease ontogeny and evolution.

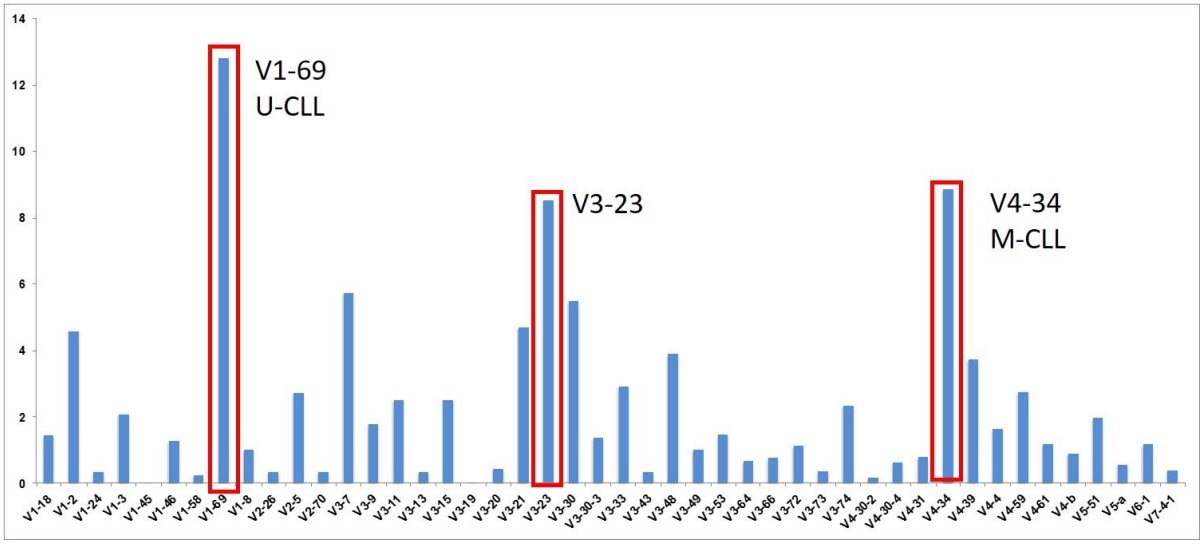


***Figure 6**: IG gene repertoire restrictions in CLL. Certain genes such as IGHV1-69, IGHV4-34 and IGHV3-23 are very frequent in CLL. A preference regarding the mutational status has also been noticed: IGHV4-34 is more frequent in M-CLL cases, whereas IGHV1-69 in U-CLL (Fais et al. J Clin Invest 1998, Agathangelidis et al. Blood 2012)*

The most striking observation to emerge from IG gene sequence data is that unrelated CLL patients carried identical, or almost identical BcR IG, a phenomenon termed as "BcR stereotypy".[99] These stereotyped BcR IGs exhibit highly homologous CDR3s that share similar amino acid motifs and, hence, similar protein structure. Notwithstanding that the likelihood of having two different clones of B cells with the same BcR by chance is extremely low ($\sim 1\times10^{-12}$), approximately 30% of CLL patients possess stereotyped BcRs.[100] BcR stereotypy is more frequent in U-CLL (>40%) rather than M-CLL (~10%).[98,101]

A large number of stereotyped subsets exist, each defined by a unique VH CDR3 motif. Agathangelidis et al studied 7596 productive IGHV-IGHD-IGHJ rearrangements from 7424 CLL patients and identified hundreds of stereotyped subsets based on shared VH CDR3 amino acid sequence patterns. Nineteen different subsets contained 20 or more

(up to 213) sequences and were defined as major.[100] (Figure 7) Even more interesting is the finding that CLL patients belonging to a distinct 'stereotyped' subset, presented remarkably similar clinicobiological features.[102] For instance, subset #2 (IGHV3-21/IGLV3-21) has been associated with poor prognosis, whereas subset #4 display a very indolent course of the disease.[98]

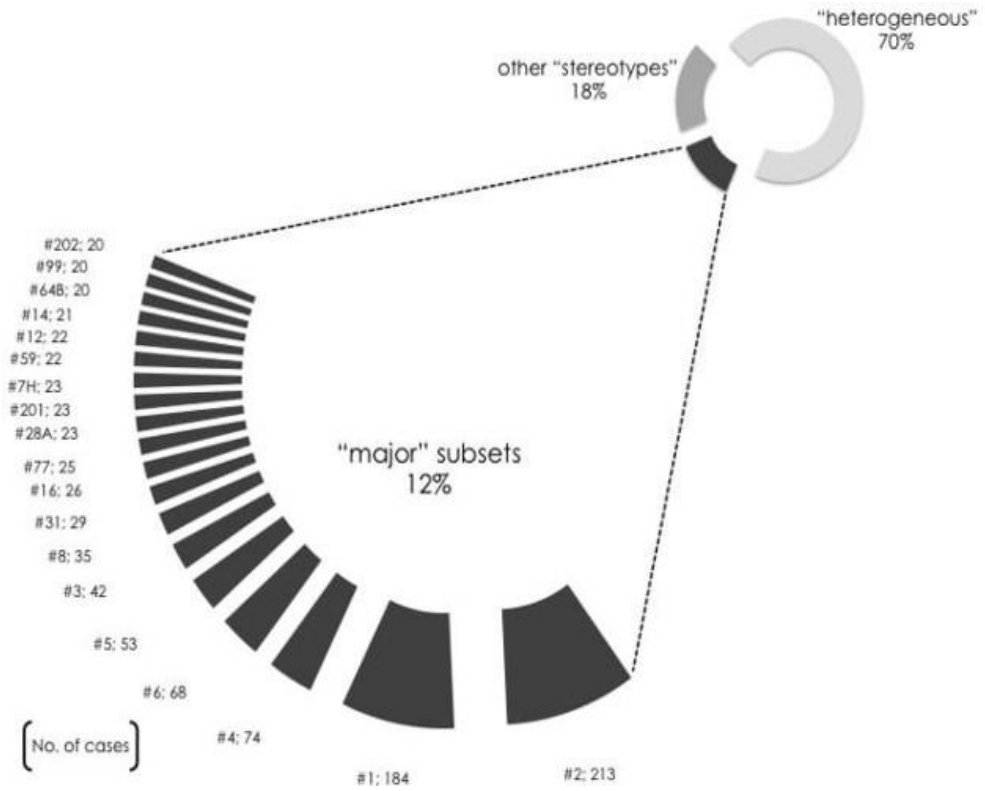


***Figure 7**: Frequency of heterogeneous and stereotyped rearrangements of immunoglobulin genes in CLL. Nineteen different subsets that contained 20 or more [up to 213 from the 7596 Ig VH (IGHV-IGHD-IGHJ) sequences analyzed] sequences were defined as major and accounted for ~41% of the stereotypes and for 12% of the cohort sequences. (Agathangelidis, A. et al, Blood, 2012)*

7.4 Antigens in CLL

Although it was initially thought that CLL cells from patients with unmutated IGHV genes originated from naïve B cells, several studies challenged this view, implicating antigens in disease ontogeny.[103] Indeed, CLL cells, independently of the SHM status of the clonotypic BcR IG, presented an immunophenotype similar to that of antigen-experienced B cells. Especially, the memory B cell marker CD27 was found to be expressed in all CLL cases, indicating that neoplastic cells had encountered antigen.[104]

Moreover, several studies, examining the BcR function, have proposed that CLL cells exhibit features of chronic antigen exposure. For example, downregulation of surface IgM but not surface IgD (a characteristic of normal anergic B cells) was observed in CLL cells.[105] In addition, U-CLL cases that do not express ZAP-70 and/or CD38, do not respond to stimulation through their surface IG, a situation that resembles that of anergic cells after chronic exposure to antigen.[105]

Today, it is generally accepted that CLL cells are antigen-experienced cells. The exact nature of antigen(s) that are recognized by the leukemic clone has become an area of intensive research. Early studies suggested autoantigenic recognition by the CLL BcR IG and these autoantigens include: the Fc (fragment crystallizable region) of IgG, ssDNA(single stranded DNA), dsDNA (double stranded DNA), histones, cardiolipin, or cytoskeletal components such as actin, tubulin, myosin.[106,107]A milestone in this research field came in 2005, when it was demonstrated that CLL antibodies from U-CLL B cells expressed highly polyreactive antibodies whereas most M-CLL B cells did not. In addition, patients expressing stereotyped BcR IG were recognizing a similar panel of antigens, implying that the stereotypy at the level of sequences strongly corresponds to specific antigen recognition.[108]

More recent studies suggested that IGs expressed in CLL recognize structures of apoptotic cells and bacteria, thus resembling natural antibodies.[109,110] On these grounds, it was proposed that CLL may derive from normal B cells whose function is to remove cellular debris, and to provide a first line of defense against pathogen.

Moreover, several cytoskeletal proteins such as vimentin, cofilin, filamin B and MYHIIA (nonmuscle myosin heavy chain IIA) can react with CLL IGs.[109,111] Apoptotic cells display MYHIIA as well as oxidized moieties on their surface membranes, which have been shown to bind IGs expressed in CLL.[112] Reactivity against autoantigens associated with systemic autoimmunity (such as Sm, Ku, snRNP and CENP-B) was also observed.[110] In a recent study, Calreticulin (a known autoantigen for ulcerative colitis, systemic lupus erythematosus and rheumatoid arthritis) was also found to function as a putative CLL antigen in specific subsets.[113]

In addition to self-antigens, several other microbial or virus-associated antigens may contribute to the selection of the B cell clones. Emerging evidence has led to the hypothesis that chronic infections by common pathogens could stimulate the precursors

of CLL cells, thus contributing to the transition from normal to neoplastic state and triggering clonal expansion of leukemic cells, at least in some subsets of patients.

A number of studies have demonstrated reactivity of CLL IGs against Gram positive or negative bacteria, such as *Streptococcus pneumoniae, Streptococcus pyogenes, Enterococcus faecium, Enterococcus faecalis* and *Enterobacter cloacae*.[109,114,115] In 2016, Hatzi et al studied the antigen-binding characteristics of 23 CLL IGs with 5 pathogenic bacteria and found that IGs differ in bacterial reactivity according to the IGHV, IGHD and IGHJ gene usage, as well as the mutational status.[116]

In addition, viruses may also be implicated in CLL pathogenesis. IGs from patients belonging to subset #4 show bias towards the IGHV4-34 gene[117], that is also known to dominate in the repertoire of otherwise healthy individuals after EBV and CMV infection.[118,119] Furthermore, results of a recent meta-analysis support an association between HBV infection and Non-Hodgkin lymphomas, including CLL.[120] In 2017, Keyvani et al found that HHV-8 DNA sequences were detected in some patients with lymphoproliferative and chronic blood disorders, indicating that HHV-8 can be considered as one of the predisposing factors for such malignancies.[121]

7.5 Autonomous signaling in CLL

The fundamental role of classical antigen stimulation in CLL was challenged when it was demonstrated that CLL BcR IGs possess antigen-independent, cell autonomous signaling capacity.[122] In particular, BcR IGs recognize themselves as autoantigens, thus initiating intracellular signaling by inducing Ca^{2+} influx and NF-κB target gene transcription. Furthermore, it was proposed that the autonomous signaling occurred through the recognition of a single, common BcR-related epitope, conserved in all cases.[123] However, in 2017 structural and biochemical analyses revealed that the autonomous signaling indeed occurs albeit with differential epitopes, intensity and functional impact in CLL cases with different prognosis.[122] The avidity of BcR self-recognition was found to directly underlie the clinical course of CLL, since tight, persistent binding was observed in cases with indolent disease, whereas weaker interactions were observed in clinically aggressive cases. [122]

8 Splenic Marginal Zone Lymphoma

8.1 Clinical and biological features

Marginal zone lymphomas (MZLs) constitute a group of mature lymphoid malignancies.[124] Three different types of MZLs are recognized by the 2016 update of the World Health Organization (WHO) classification of tumors of the hematopoietic and lymphoid tissues:

1) Nodal Marginal Zone Lymphoma (NMZL)
2) Extranodal Marginal Zone Lymphoma of mucosa-associated lymphoid tissue (MALT)
3) Splenic Marginal Zone Lymphoma (SMZL)[124]

SMZL is a distinct indolent B cell lymphoma that predominantly affects elderly patients and involves the spleen, bone marrow and (frequently) peripheral blood but not peripheral lymph nodes. The median age at diagnosis is 69 years and the median survival is estimated at around 10 years.[125,126] SMZL is a rare neoplasm representing 20% of MZL and 0.6% of non-Hodgkin lymphoma (NHL) cases in the Surveillance, Epidemiology, and End Results registries.[127] Moreover, according to a population-based study for the United States, the overall age-adjusted incidence is 0.13 per 100 000 persons per year with an increased incidence among people who are white or male.[127]

Studies support an association between HCV infection and SMZL.[128] In addition, SMZL is characterized by high prevalence of autoimmune phenomena such as hemolytic anemia, immune thrombocytopenia, acquired coagulation disorders and positive Coomb's test.[129]

Although a specific immunophenotype is not evident for SMZL, in most SMZL patients, the tumoral monoclonal B cells express surface IgM and IgD or IgM alone. Moreover, they are positive for CD20, CD22, CD24, CD27, FMC7, Pax5, BCL2 and display strong expression of CD79b.[130] DBA44 (CD76), CD11c, CD23, CD103 and CD25 may be present in some cases, whereas tumor cells typically do not express CD3, CD5, CD10, CD23, CD43, cyclin D1, annexin-A1 or BCL6.[126]

A considerable number of patients exhibit cytogenetic abnormalities, however SMZL lacks recurrent chromosomal translocations, including translocations that are typical of other lymphomas.[131] In 2010, Salido et al, described in detail the cytogenetic findings in a large group of well-characterized SMZLs, where of 330 cases, 72% had an aberrant karyotype. Within this subgroup, 53% displayed a complex karyotype (3 or more cytogenetic abnormalities).[132] Deletion at the 7q21-q36 band, seems to be more frequent in SMZL compared to other lymphoma subtypes. However, the genes affected by this deletion still remain to be elucidated.

Moreover, genes of the NOTCH pathway are mutated in 40% of the patients.[133] *NOTCH2* is one of the most frequently mutated genes in SMZL and may represent a diagnostic and prognostic tool, since it is highly specific for this lymphoma.[134] Furthermore, *TP53* mutations have been observed in 16% of patients and identified as independent markers of poor outcome.[134] Krüppel-like Factor 2 (KLF2) protein, a transcription factor that regulates the expression of B cell genes, is transcriptionally defective, in 20% to 40% of SMZL cases.[126,135] *KLF2* is one of the most frequently altered genes in SMZL and has been associated with poor outcome.[126,134]

Studies have also revealed mutations in genes involved in chromatin remodeling. *MLL2*, encoding a histone methyltransferase, is inactivated in 15% of SMZL cases, *EP300* harbors somatic missense mutations in 5% of SMZL, whereas *SIN3A*, a crucial component of the SIN3–HDAC1/2 histone deacetylase complex carries mutations in 7.5% of SMZL cases.[136] *ARID1A*, a member of the SWI-SNF chromatin remodeling family, the *CREBBP* acetyltransferase and other chromatin remodeling genes are also affected.[136]

8.2 Immunogenetic findings in SMZL

In 2012, Bikos et al. examined the IGHV-IGHD-IGHJ rearrangements from 337 patients and showed that only three IGHV genes namely IGHV1-2, IGHV4-34 and IGHV3-23 accounted for 45.8% of the cases.[137] Moreover, the IGHV1-2*04 allele predominated by far amongst all genes (~30% of all SMZL), strongly alluding to functional selection.[137] Intriguingly, IGHV1-2*04 rearrangements carried long VH CDR3s and showed biased recombination with one of the IGHD3-3, IGHD3-9 or IGHD3-10 genes, generating distinct motifs, which may allude to the recognition of

specific antigenic epitopes.[137] Additionally, 95% of IGHV1-2*04 rearrangements carried SHMs and although these were relatively few their distribution was quite restricted, often affecting FR codons.[137]

Subsequent analysis of the IG LC gene repertoire in 107 SMZL patients also revealed biases with six genes accounting for 64% of the cases, namely IGKV3-20, IGKV4-1, IGKV1-5, IGKV1-8, IGKV1-39 and IGLV2-14.[138] Furthermore, the IGHV1-2*04 cases with available IG LC sequence exhibited biased usage of three LC genes (IGKV3-20, IGKV1-8, IGLV2-14), further supporting the notion that antigen selection could be involved in SMZL ontogeny. (Figure 8 and 9)

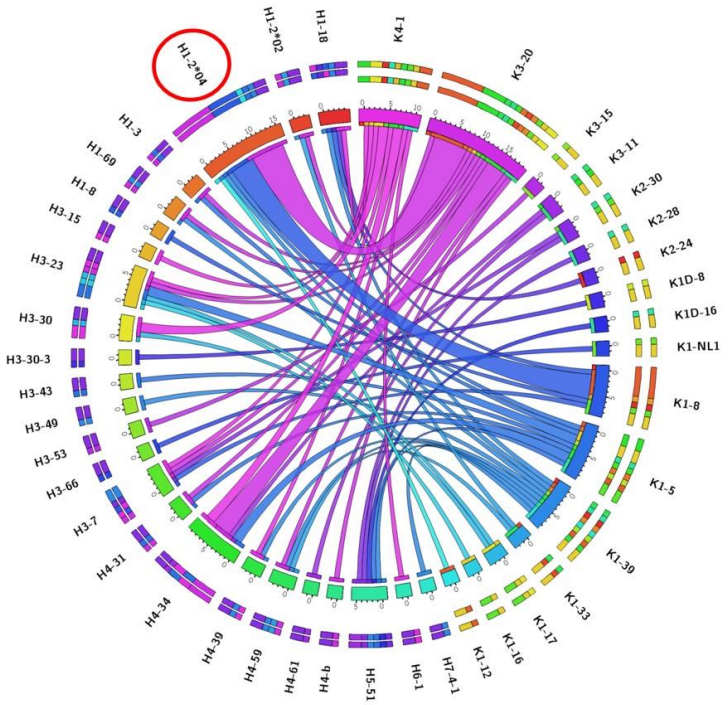


***Figure 8**: Associations of IGHV genes with IGKV genes. Biased associations of the IGHV1-2*04 gene with the IGKV3-20 and IGKV3-15 genes. (Bikos, V. et al. Leukemia, 2012)*

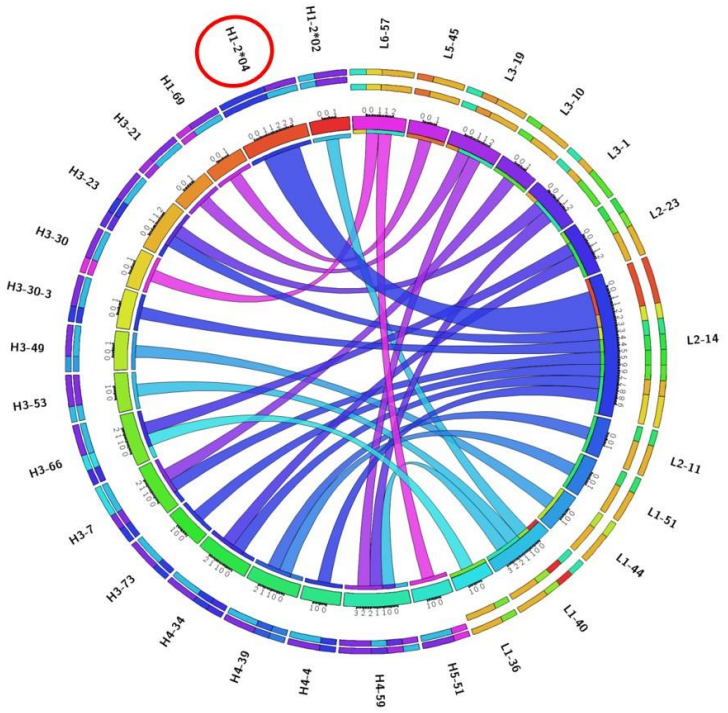


***Figure 9**: Associations of IGHV genes with IGLV genes. Biased association of the IGHV1-2*04 gene with the IGLV2-14 gene. (Bikos, V. et al. Leukemia, 2012)*

It is also worth mentioning that examination of sequence variation in the IGHV and IGK/LV genes has revealed extensive and distinctive Intraclonal Diversification (ID), suggestive of clonal evolution.[139] This finding supports the idea that antigens are involved not only in selecting the B cell progenitors but also post-transformation by applying an ongoing pressure on the tumor cells. In addition, ID is more frequent in IGHV1-2*04 SMZL, implying that these cases might constitute a distinct molecular subtype of the disease, with implications for both SMZL subclassification and research.[139]

8.3 Antigen stimulation in SMZL pathogenesis

In SMZL, restrictions in IG gene usage suggest that a limited number of antigens may be involved in the disease ontogeny and evolution. Published evidence implicates chronic antigen stimulation by the Hepatitis C virus, the malaria parasite, or an autoantigen in the ontogeny of at least some SMZL cases.[140–143] Initially, this stimulation may result in polyclonal expansion of B cells in the splenic MZ, where subsequent deregulation of genes such as *TP53, KLF2* and *NOTCH2* may trigger the outgrowth of a monoclonal population.

Given their high frequency, IGHV1-2*04 expressing cases have attracted a lot of interest. Recombinant monoclonal antibodies (mAbs) produced from patients carrying this gene were reported to react with numerous molecules, such as nuclear, cytoplasmic and membranous antigens expressed by human cells as well as human serum.[144] However, these mAbs do not react with proteins on apoptotic cells, in contrast to what has been demonstrated in CLL.[145] Overall, on these grounds, it was suggested that IGHV1-2*04 SMZL cases may arise from polyreactive B cells.

Aim of the study

Mounting evidence suggests that microenvironmental triggering by various cells and soluble factors, including antigens, is critical for the development and evolution of several malignancies of mature B cells, including CLL and SMZL. Indeed, the crucial role of antigenic stimulation in B lymphomagenesis is underscored by the restrictions in the IG gene repertoire of clonotypic BcRs and the characteristic imprints of SHM within the BcR IG genes, in certain cases leading to intraclonal diversification ascribed to continuous interaction with antigen(s).

In CLL, perhaps the strongest argument favoring antigen selection concerns BcR IG stereotypy i.e. the fact that unrelated patients carry (quasi)identical BcR IG.[99] Moreover, considering that CLL patients belonging to a distinct 'stereotyped' subset display remarkably similar clinicobiological features, it is not unreasonable to speculate that a particular antigenic element binding to a distinctive receptor may be critical in modulating the natural history of the disease and determining clinical presentation.[102] Of note, BcR IG restrictions are also present in SMZL, albeit of a different kind than in CLL, supporting antigen selection in this disease as well.[137,139]

Recent studies have shown that CLL BcR IGs resemble natural antibodies recognizing molecular motifs on both apoptotic cells (e.g. modified cytoskeletal proteins and oxidation-specific epitopes) as well as exogenous microorganisms, indicating that CLL clones may arise from B cells acting as scavengers of apoptotic debris, while also being able to bind conserved microbial structures. However, despite the progress achieved, there are still major unresolved issues especially regarding the nature of the implicated antigens and their precise functional impact on clonal evolution. On the other hand, for SMZL, the single study published thus far concerned IGHV1-2*04 expressing clones and showed poly- and self-reactivity with recognition of nuclear, cytoplasmic and cell membrane antigens and human serum.[144]

The aim of this study was to identify antigen(s) recognized by the clonotypic BcR IGs deriving from CLL and SMZL cases. This has the potential to contribute to: (i) the elucidation of the molecular mechanisms underlying CLL and SMZL pathogenesis, (ii) the identification of a possible causal relationship between CLL and SMZL with certain pathogens, (iii) the redefinition of possible risk factors and epidemiology of the

respective disease, (iv) the classification of patients according to the type of antigenic stimulation, (v) the development of targeted treatment approaches by interfering in the signaling between the antigen and leukemic clone.

Materials and methods

1. Recombinant monoclonal IG production

1.1 Plasmid vectors

The pIgMu, pIgKappa and pIgLambda plasmids were used for cloning. (Figure 10, 11 and 12, respectively) These plasmids encode the constant region of human IgM heavy chain and human Kappa and Lambda light chains, respectively, providing a framework to insert the variable regions of each IG chain. Upstream of the constant region sequence, there is a multiple cloning site (MCS) containing a range of standard commonly used restriction sites for cloning and a murine leader peptide sequence (GenBank accession no. DQ407610).

Expression is based on the Human Cytomegalovirus Promoter (HCMV) leading to efficient transcription of the gene, whereas replication is done through the pUC, SV40 and F1ori. Moreover, the expression constructs include a gene that confers resistance to ampicillin allowing only cells containing the plasmid to survive in growth medium having the particular antibiotic. IG variable regions were cloned by PCR, between AgeI and AfeI restriction sites for plasmid pIgMu, AgeI and BsiWI restriction sites for plasmid pIgKappa and AgeI and XhoI restriction sites for plasmid pIgLambda.

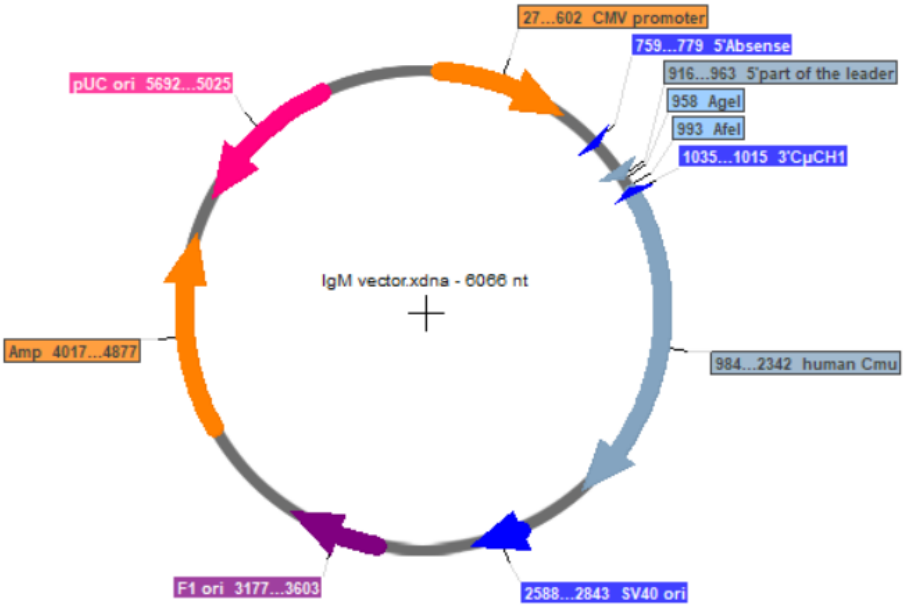


***Figure 10**: Map of the vector pIgMu for mAb production.*

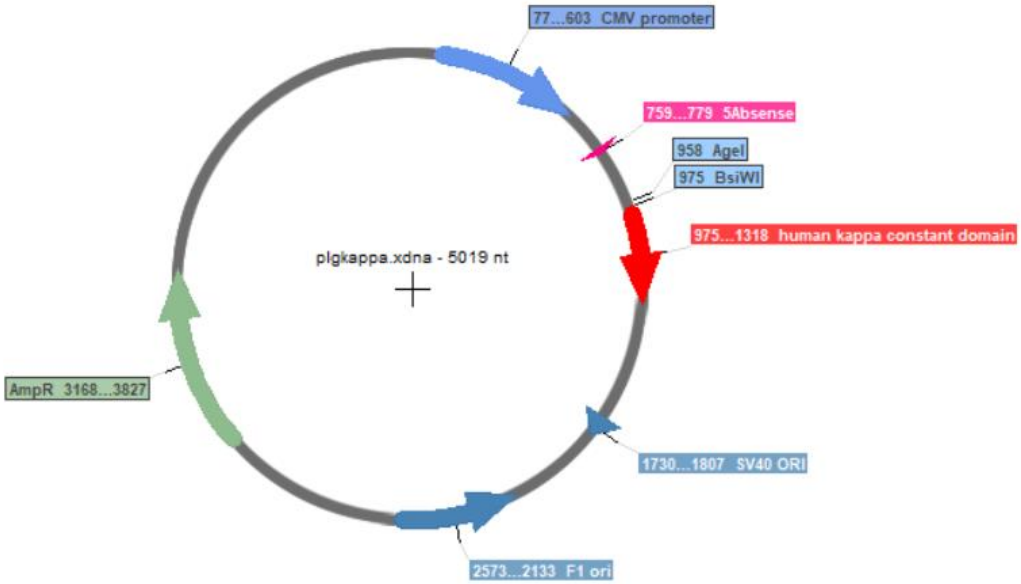


***Figure 11**: Map of the vector pIgKappa for mAb production.*

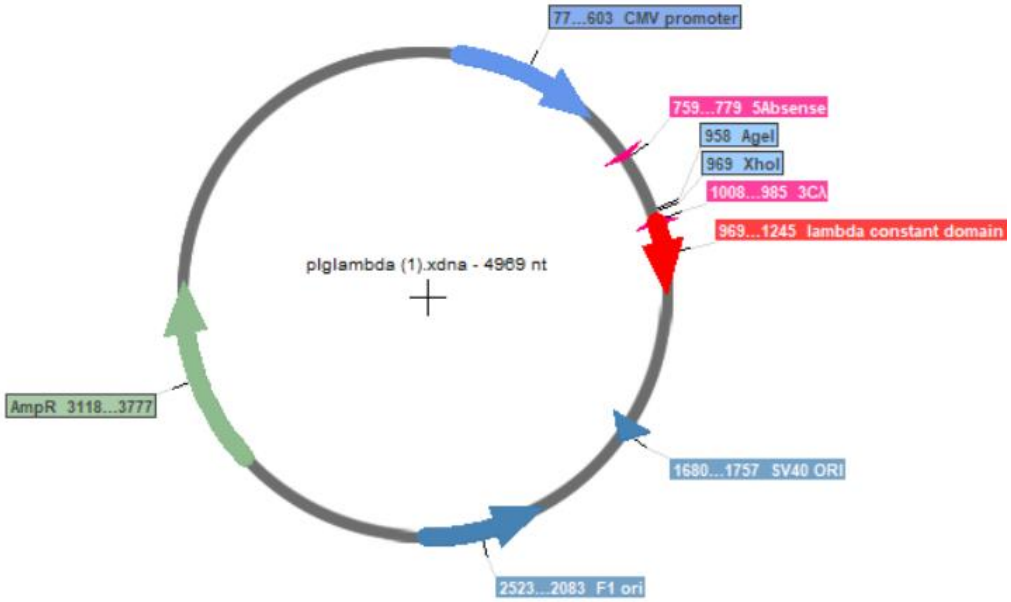


***Figure 12**: Map of the vector pIgLambda for mAb production.*

1.2 Cloning by the PCR technique

The PCR method is widely used for cloning DNA into various plasmid vectors.[146] In the present study, cDNA of the VH and VK or VL domains was cloned in the plasmid vectors pIgMu, pIgKappa and pIgLambda, by PCR.

Since all cases of the present study express clonotypic BcR IG of the mu isotype, all monoclonal antibodies were produced as human monoclonal IgMs.[147,148] In addition, pIgKappa or pIgLambda plasmids were used for cloning, based on the light chain that was expressed by each leukemic clone.

First, for the identification of the IGHV-IGHD-IGHJ, IGKV-IGKJ and IGLV-IGLJ rearrangements of each leukemic clone studied, PCR was performed as described.[149] Next, PCR products were purified with the Monarch® DNA Gel Extraction Kit (NEB) according to the manufacturer's instructions and sequenced with the appropriate primer. Sequences were analyzed using the IMGT/V-QUEST tool (http://www.imgt.org/IMGT_vquest/vquest) in order to identify the germline V(D)J genes with highest identity and determine the CDR3 length and composition.

Subsequently, gene-specific PCR was performed using appropriate oligonucleotide primers for the amplification of the IG variable domain and the insertion of restriction sites for AgeI, AfeI, BsiWI and XhoI restriction enzymes.[149]

The reagents for the amplification of IGHV-IGHD-IGHJ and IGKV-IGKJ/IGLV-IGLJ rearrangements and the PCR conditions are presented in Tables 1 and 2, respectively.

Table 1: *Reagents for the amplification of IGHV-IGHD-IGHJ and IGKV-IGKJ/IGLV-IGLJ rearrangements*

Reagents	Volume (μl)
10X Reaction Buffer	5
dNTP mix	2,5
Forward primer	1,5
Reverse primer	1,5
H_2O	37,5
Pfu DNA Polymerase	1
Total volume	50

Table 2: *Thermocycler conditions for cloning*

Thermocycler conditions		
Initial denaturation	94ºC 5min	
Denaturation	94ºC 30sec	x45
Annealing	58ºC for IgH/Igκ or 60ºC for Igλ, 30sec	
Extension	72ºC 30sec	
Final extension	72ºC 5min	
Storage	4ºC	

Digests of PCR products were carried out with the appropriate restriction enzymes (all from NEB, Table 3): AgeI and AfeI for the mu heavy chain, AgeI and BsiWi for the Kappa light chain and AgeI and XhoI for the Lambda light chain. The same restriction enzymes were used for digestion of the corresponding vectors. Digests were performed simultaneously or sequentially depending on the requirements of each enzyme. Ten to 20 units of enzyme and 0.1 to 2 µg DNA were used. The incubation time ranged from 1.5 to 3 hours with a typical time of 2 hours. Digested products were purified as described before.[146,149]

Table 3*: Description of restriction enzymes used for the cloning of IG variable domains*

Enzyme	Restriction site	Temperature
AgeI	A/CCGGT TGGCG/A	37ºC
AfeI	AGC/GCT TCG/CGA	37 ºC
BsiWi	C/GTACG GCATG/C	55 ºC
XhoI	C/TCGAG GAGCT/C	37 ºC

1.3 Ligation

Ligation was performed in a total volume of 10 µl with 10 units T4-Ligase (Invitrogen), 100-200 ng digested plasmid vector, three times excess molar quantity of digested PCR product and appropriate buffer solution (Invitrogen) at a final concentration 1x. The mixture was incubated at 14 °C for 12-16 hours.

1.4 Transformation of bacteria with recombinant plasmids

Chemically competent Top10 bacteria (Invitrogen) were transformed with 5 μl of the ligation product. Colonies were screened by PCR using as forward primer 5′Absense, annealing upstream the insert, and 3′ CmCH1, 3′Cκ494 or 3′Cλ as reverse primer for the mu heavy chain, kappa light chain and lambda light chain plasmids, respectively, that anneal just downstream the insert. The sequences of the primers and the PCR conditions are given in Tables 4 and 5.[149]

Colonies with PCR products of the expected size (~650bp for IgM, 700 bp for Igκ and 590 bp for Igλ) were selected for amplification in 20 ml bacteria cultures and in LB Broth Base (Lennox L Broth Base- Invitrogen) containing 100 μg/ml ampicillin (Sigma). Next, plasmid DNA was isolated, using QIAprep Spin columns (Qiagen). Subsequently, plasmid DNA was sequenced and analyzed using the IMGT/V-QUEST tool, to confirm identity with the original PCR products.

***Table 4**: Primer sequences for the confirmation of the insertion of the IG genes into the plasmid vectors*

Primer	5′ - 3′ sequence	Vector
5’ Absense	GCTTCGTTAGAACGCGGCTAC	pIgM, pIgκ, pIgλ
3′ CmCH1	GGGAATTCTCACAGGAGACGA	pIgM
3'Ck494	GTGCTGTCCTTGCTGTCCTGCT	pIgk
3'Clamb	CACCAGTGTGGCCTTGTTGGC	pIgλ

***Table 5**: Thermocycler conditions for bacterial colony screening*

Thermocycler conditions		
Initial denaturation	94°C 5min	
Denaturation	94°C 30sec	X35
Annealing	58°C, 30sec	
Extension	72°C 1min	
Final extension	72°C 10min	
Storage	4°C	

2. Production of recombinant monoclonal IGs

2.1 Human embryonic kidney (HEK) 293T cell cultures

Human embryonic kidney (HEK) 293T cells were cultured in 100 mm plates, for adherent cells, in Dulbecco's Modified Eagle's Medium (DMEM; GibcoBRL) supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS) and 0,1% (v/v) penicillin and streptomycin. Cells were incubated at 37°C in 5% CO_2 and when they reached 70-90% confluency, they were objected to split procedure as follows:

- aspiration of the medium and washing with 10 ml PBS 1x
- addition of 1 ml trypsin for a couple of minutes (depending on the responsiveness of the cells)
- dilution with 10 volumes (10 ml) of medium containing FBS, gently pipetting several times to ensure an even suspension
- transfer to new plates by supplementing with fresh medium.

The dilution of the initial culture ranged from 1: 4 to 1: 8.

2.2 Transfection of HEK 293T cells

HEK293T cells were split one day before transfection in order to have a confluency of 50-70% at the time of transfection. For each case, four plates were prepared. On the day of the transfection, a few hours before the procedure, DMEM (supplemented with 10% FBS and 0,1% penicillin-streptomycin) was replaced with fresh one.

Transient transfection of cells was performed by using Polyethylenimine (PEI) at a 3:1 ratio of PEI to DNA as follows:

- mix a dilution of 5 μg plasmid DNA containing the heavy chain genes and 5 μg of plasmid DNA containing the light chain genes in 0,5 ml Optimem
- in another sterile tube, add a dilution of 30 μg PEI in 0,5 ml Optimem
- transfer the diluted PEI to the diluted DNA with immediate vortexing or pipetting. incubation of the mixture for 15 minutes, at room temperature
- transfer of 1 ml of the PEI-DNA mixture to each plate drop-wise on top of the medium.

Twelve to sixteen hours later, medium was replaced with DMEM supplemented with antibiotics and 1x Nutridoma-SP (Roche). Nutridoma-SP was used instead of FBS in

order to facilitate the subsequent purification of monoclonal IGs from the culture supernatant.

2.3 Collection and concentration of recombinant monoclonal IGs

Culture supernatants were collected after 6-7 days of cell culture, cell debris was removed by centrifugation at 800×g for 10 minutes and supernatants were subjected to ~25-fold concentration, using Millipore® Amicon® Ultra-15 Centifugal Filter Concentrators with a Molecular Weight Cutoff 100kDa. Concentrated supernatants were stored at 4°C with 0,05% sodium azide and protease inhibitors.

2.4 Quantification of recombinant monoclonal IGs

Recombinant IG concentrations were determined by Human IgM ELISA development kits (Mabtech), following the manufacturer's instructions. Briefly, high-binding capacity ELISA plates (Costar) were coated with 100 μl capture anti-IgM antibodies at a concentration of 2 μg/ml in PBS. Plates were washed twice with 200 μl of PBS per well, before blocking for 1 hour with 200 μl PBS, 0,05% Tween-20, containing 0,1%BSA and washed again five times. Standards and samples diluted appropriately were added and were incubated for 2 hours at room temperature. Unbound IGs were removed by washing 5 times, before incubation with 100 μl/well of alkaline phosphatase (ALP)-coupled anti-human IgM for 1 hour at room temperature. Unbound secondary antibody was removed by washing as previously. Assays were developed using 50 μl of the colorimetric, soluble p-nitrophenyl phosphate (pNPP) substrate and optical density (OD) was measured at 405 nm using the ELx800™ Absorbance Microplate Reader (BioTek).

2.5 SDS-PAGE of recombinant monoclonal IGs

The purity and the integrity of the monoclonal IGs were examined by polyacrylamide gel electrophoresis, using the Invitrogen X Cell SureLock device and NuPAGE® 4-20% Tris-Glycine Gels. Electrophoresis was performed under denaturating and non-denaturating conditions, at 200 V for 2 hours in Tris Glycine 1% SDS running buffer.

Markers of known molecular weight were also used (Blue Star pre-stain marker and MagicMark XP, Invitrogen).

Following SDS-PAGE, proteins were transferred to a PVDF membrane (BioRad, USA) with Tris-Glycine 1x transfer buffer, containing 10% methanol and by applying 35 V, for 2 hours. Successful transfer was confirmed with Ponceau S solution and destaining was performed with 2-3 washes with PBS containing 0,001% Tween (PBS-Tween) for 5 minutes.

Subsequently, membrane was blocked with 5% skim milk in PBS-T for 1 hour, at room temperature. Then, Horseradish Peroxidase (HRP)-conjugated goat anti-Human IgM (Abcam) diluted 1:20.000 in 5% milk was added for 40 minutes, at room temperature. Next, the membrane was washed 3 times with PBS-T and the proteins were detected by the enhanced chemiluminescence method.

3. Antigen reactivity examination by the ELISA technique

3.1 Anti-Nuclear Antibody reactivity

The produced recombinant mAbs were tested for Anti-Nuclear Antibody (ANA) activity, using precoated plates, containing ANA antigens wells, provided by a commercially available ELISA kit (Hycor Biomedical). Recombinant CLL or SMZL IgM mAbs were added as primary antibodies at a concentration of 25 μg/ml. After 1,5 hour incubation, the ELISA wells were washed 3 times and 100 μl of Anti-Human IgM HRP conjugated (abcam) was added as secondary antibody for 1.5 hour. The wells were washed again and the assay was developed using 100 μl of Peroxidase Substrate (3,3',5,5' tetramethylbenzidine, TMB). After 15 minutes incubation in the dark, the reaction was stopped by adding 0.18M H_2SO_4 (stop solution) and the OD was measured at 450 nm using the ELx800™ Absorbance Microplate Reader (BioTek).

All recombinant mAbs were examined twice and all steps were conducted at ambient temperature.

3.2 Anti double-stranded DNA reactivity

Recombinant mAbs were examined for reactivity against double-stranded DNA (dsDNA) by ELISA. Briefly, high-binding capacity ELISA plates (Costar) were coated with 100μl deoxyribonucleic acid from calf thymus (Sigma Aldrich) at a concentration of 5 μg/ml in PBS. Plates were washed 5 times with PBS, 0,1%Tween, blocked for 1 hour with PBS containing 1,5% BSA and washed again 5 times. CLL and SMZL mAbs were added at various concentrations (0,5, 5 and 10 μg/ml) at the ELISA plates. After overnight-incubation at 4°C, unbound IGs were removed by washing 5 times and anti-Human IgM HRP conjugated (abcam) was added in 1:20000 dilution for 2 hours, at room temperature. The plate was washed 5 times and the assay was developed using 50 μl of TMB Peroxidase Substrate (Thermo Scientific) with incubation for 15-30 minutes. The reaction was stopped by adding 50μl of TMB stop solution and OD was measured at 450 nm using the ELx800™ Absorbance Microplate Reader (BioTek).

3.3 Evaluation of the ELISA results

Evaluation of antigen recognition was based on a scale utilized in previous ELISA experiments.[110] The ratio of sample's OD to the OD of the negative control (background) was calculated and the results were interpreted as follows: - = <2.5, + = 2.5-5, + + = 5 - 10, + + + = 10-15, + + + + => 15.

4. Flow cytometry

25x10^{4} MEC1 cells were resuspended at 100 μl PBS-FBS 1% and incubated with 25 μg/ml recombinant SMZL or CLL mAbs, for 1 hour at RT. Then, they were washed with PBS-FBS 1%, incubated with 10 μl IgM κ-FITC/λ-PE secondary Ab (BD Biosciences) for 15 minutes in the dark and washed again in PBS. After resuspension in 500 μl PBS, 4 μl of 7AAD (viability dye) was added and data were collected using a BD FACS Calibur flow and analyzed with Flowing software (http://flowingsoftware.btk.fi/).

5. Western blot analysis in MEC1 cell extracts

To examine if the produced recombinant mAbs recognize cytoplasmic or membrane protein antigens of MEC1 cells, whole cell extracts were isolated from $5x10^6$ MEC1 cells, using 100 µl of cold lysis solution, containing RIPA buffer, protease inhibitors (1/10) and benzonase (1/1000).

Next, 20 µg MEC1 were loaded on each lane of a Tris-Glycine gel. Electrophoresis and transfer of the proteins to PVDF membrane was performed as described above.

Subsequently, the membrane was blocked with 5% skim milk in PBS-T for 1 hour, at room temperature. Each lane of the membrane, all loaded with MEC1 lysate, was cut in order to get blotted with a different recombinant mAb. In total, 6 different recombinant mAbs were used as primary antibodies, diluted 1:20000 in 5% skim milk in PBS-T and incubated overnight at 4ºC. Membranes were washed 3 times with PBS-T and then anti-Human IgM HRP conjugated (abcam) or anti-Human IgG (abcam) diluted 1:20.000 in 5% milk was added for 1 hour, at room temperature. Subsequently, they were washed 3 times with PBS-T and immunodetection was performed by the enhanced chemiluminescence method.

6. Statistical analysis

Statistical analysis and graphical representation was carried out using Microsoft Excel 2016 and Graphpad Prism 5.0 software (La Jolla, CA, USA). For quantitative data, average and standard deviation were calculated. Comparative analyzes were performed using t-test for independent samples (non-parametric Mann-Whitney test). All analyzes were made at a 5% significance level.

Results

1. Production of recombinant monoclonal antibodies (mAbs)

Plasmids containing the variable regions of both heavy and light IG chains from 8 SMZL and 34 CLL patients were produced. Next, the corresponding mAbs were prepared by transient transfection of HEK293T cells as recombinant human IgM paired with kappa or lambda light chain, depending on the light chain isotype expressed by the malignant clone of each patient.

The IG gene repertoire of the cases included in the study group, as well and the grouping of the CLL cases into stereotyped subsets, are given in Table 6.

In frame cloning of the variable IG regions with the respective human Igμ, Igκ or Igλ constant region genes encoded by the eukaryotic expression vectors was confirmed by sequencing analysis. (Tables 6-7, Figure 13)

***Table 6**: . IGHV-IGHD-IGHJ sequences of recombinant mAbs included in the study. CLL stereotyped subset nomenclature according to Agathangelidis et al, 2012 and Hoogeboom, R et al. 2013.*

Case ID	Isotype	Entity and S.S.	IGHV	% gene identity	IGHD	IGHJ	CDR3 amino acid sequence
P8761	μ/κ	SMZL	IGHV1-2*04	100.00	IGHD3-3*01	IGHJ5*02	CARGGRASVFGVVLIGDPIFGPW
P19994	μ/κ	SMZL	IGHV1-2*04	100.00	IGHD3-22*01	IGHJ4*02	CARDQGGGSGYYHYFDYW
P411	μ/κ	SMZL	IGHV1-2*04	97.57	IGHD3-3*01	IGHJ5*02	CARGGRITIFGVLMGRENWFDPW
P23776	μ/λ	SMZL	IGHV1-2*04	98.96	IGHD3-3*01	IGHJ5*02	CARGPRITIFGVVIGRGTTLDPW
P13836	μ/λ	SMZL	IGHV1-2*04	94.10	IGHD3-10*01	IGHJ4*02	CAKDRARTYYYGSGSYYNRPPQYYFDYW
P2120	μ/κ	SMZL	IGHV3-23*01	97.92	IGHD4-17*01	IGHJ4*02	CAKKGDQAPYGDNLYYFDYW
P8257	μ/κ	SMZL	IGHV4-34*01	100.00	IGHD2-21*01	IGHJ6*03	CARVWGGLVGSGYFYSGYYYYYMDVW
P12561	μ/κ	SMZL	IGHV3-30-3*01	94.44	IGHD4-17*01	IGHJ4*02	CARDWYGDYPHYFDYW
RS038	μ/κ	CLL #1	IGHV1-3*01	100.00	IGHD6-19*01	IGHJ4*02	CAREQWLGPYYFDYW
GM056	μ/κ	CLL #1	IGHV1-3*01	100.00	IGHD6-19*01	IGHJ4*02	CAREQWLAITHFDYW
P3870	μ/κ	CLL #1	IGHV1-2*02	100.00	IGHD6-19*01	IGHJ4*02	CARGQWLVQLNFDYW
P14197	μ/κ	CLL #1	IGHV1-2*02	100.00	IGHD6-19*01	IGHJ4*02	CAREQWLVRVHFDYW
P23728	μ/κ	CLL #1	IGHV1-2*02	100.00	IGHD6-19*01	IGHJ4*02	CAREQWLVLRNFDYW
P5220	μ/κ	CLL #1	IGHV1-2*02	100.00	IGHD6-19*01	IGHJ4*02	CARAQWLVLPHFDYW
P1173	μ/κ	CLL #1	IGHV5-10-1*03	99.31	IGHD6-19*01	IGHJ4*02	CAREQWLGIKNFDYW
N4755	μ/λ	CLL #2	IGHV3-21*01	97.92	IGHD5-12*01	IGHJ6*02	CATDRNGMDVW
P11475	μ/λ	CLL #2	IGHV3-21*01	97.99	ND	IGHJ6*02	CARDQNAMDVW
P326	μ/λ	CLL #2	IGHV3-21*02	97.92	IGHD3-16*01	IGHJ6*02	CAIDRNGMDVW
CLL 412	μ/λ	CLL #2	IGHV3-21*01	98.25	ND	IGHJ6*02	CARDQNGMDVW
KM1PL	μ/κ	CLL #6	IGHV1-69*06	100	IGHD3-16*02	IGHJ3*02	CARGGNYDYIWGSYRSNDAFDIW
P7150	μ/κ	CLL #6	IGHV1-69*01	100	IGHD3-16*02	IGHJ3*02	CARGGSYDYVWGSYRPNDAFDIW
P10824	μ/κ	CLL #6	IGHV1-69*01	100	IGHD3-16*02	IGHJ3*02	CARGGGYDYIWGSYRPNDAFDIW
P11887	μ/κ	CLL #6	IGHV1-69*06	100	IGHD3-16*02	IGHJ3*02	CARGGPYDYVWGSYRPSDAFDIW
P1430	μ/λ	CLL #7	IGHV1-69*01	100	IGHD3-3*01	IGHJ6*02	CAREGNYDFWSGYYPNYYYYGMDVW
P6275	μ/λ	CLL #7	IGHV1-69*01	100	IGHD3-3*01	IGHJ6*02	CATINYDFWSGYYKNYYYGMDVW
P104	μ/λ	CLL #10	IGHV4-39*07	98.97	IGHD2-2*01	IGHJ6*02	CARDRGGYCSSTSCYLFYYGMDVW
P5071	μ/λ	CLL #10	IGHV4-39*01	100	IGHD2-2*01	IGHJ6*02	CARHRLGYCSSTSCYYYYYGMDVW
BRNO_1520	μ/κ	CLL #13	IGHV4-59*01	91.93	IGHD2-15*01	IGHJ2*01	CARDSYCTGGSCFDWYFDLW
BRNO_0376	μ/κ	CLL #13	IGHV4-59*01	93.68	IGHD2-15*01	IGHJ2*01	CARDRYCSGGSCFDWYFDLW
BE064	μ/κ	CLL #59	IGHV1-58*01	100	IGHD3-3*01	IGHJ4*02	CAAGPDFWSGYPYW
P5623	μ/λ	CLL #77	IGHV4-59*03	91.93	IGHD2-21*02	IGHJ1*01	CARGPNESGWLGLLHW
P8407	μ/λ	CLL #77	IGHV4-4*02	92.01	IGHD6-19*01	IGHJ6*02	CTRGPDESGWNGMDVW
P14438	μ/λ	CLL #77	IGHV4-59*01	89.47	IGHD6-19*01	IGHJ4*02	CARGPDESGWNGLLYW
P2959	μ/λ	CLL #77	IGHV4-4*02	95.09	IGHD6-19*01	IGHJ4*02	CARGPDISGWNGFEYW
P6360	μ/κ	CLL #148B	IGHV2-5*02	94.85	IGHD3-3*01	IGHJ4*02	CARRGEDLTGWMNAYFTFW
P6540	μ/λ	CLL #169	IGHV3-48*02	96.18	IGHD1-26*01	IGHJ3*01	CARDGVGAPLW
P775	μ/λ	CLL #201	IGHV4-34*01	94.74	IGHD1-26*01	IGHJ3*01	CARRPERWDLYKNDAFDVW
L3961	μ/λ	CLL #201	IGHV4-34*01	91.23	IGHD3-10*01	IGHJ3*02	CARRPSNWELFRMDAFDIW
L7911	μ/λ	CLL #201	IGHV4-34*01	94.39	IGHD1-26*01	IGHJ3*01	CARRRERWDQKEGDAFDVW
P22229	μ/λ	CLL #201	IGHV4-34*01	91.58	IGHD3-9*01	IGHJ3*01	CARREETWTEIKGDAFDVW
fungi_725	μ/κ	CLL #fungi[150]	IGHV3-7*01	95.49	IGHD3-10*01	IGHJ5*02	CAEGTVGW
fungi_3687	μ/κ	CLL #fungi[150]	IGHV3-7*01	93.40	IGHD5-24*01	IGHJ4*02	CIESSTNW

S.S.: CLL stereotyped subset

***Table 7**: IGKV-IGKJ, IGLV-IGLJ sequences of recombinant mAbs included in the study. CLL stereotyped subset nomenclature according to Agathangelidis et al, 2012 and Hoogeboom, R et al. 2013.*

Case ID	Isotype	Entity and S.S.	IGKV/IGLV	% gene identity	IGKJ/ IGLJ	CDR3 amino acid sequence
P8761	μ/κ	SMZL	IGKV1-8*01	100.00	IGKJ1*01	CQQYYSYPRTF
P19994	μ/κ	SMZL	IGKV1-39*01	100.00	IGKJ2*01	CQQSYSTPPDTF
P411	μ/κ	SMZL	IGKV3-20*01	97.52	IGKJ3*01	CQQYGNSPPITF
P23776	μ/λ	SMZL	IGLV1-40*01	98.26	IGLJ3*02	CQSYDSSLSGSVF
P13836	μ/λ	SMZL	IGLV2-11*01	96.18	IGLJ2*01	CCSYAGFYTSKAF
P2120	μ/κ	SMZL	IGKV4-1*01	99.66	IGKJ3*01	CQQYYSTPFTF
P8257	μ/κ	SMZL	IGKV3-11*01	100.00	IGKJ2*04	CQQRSNWPPVCSF
P12561	μ/κ	SMZL	IGKV3-20*01	97.87	IGKJ1*01	CQQYGNSPGTF
RS038	μ/κ	CLL #1	IGKV1-39*01	100	IGKJ3*01	CQQSYSTPPLFTF
GM056	μ/κ	CLL #1	IGKV1-39*01	100.00	IGKJ2*01	CQQSYSTPGYTF
P3870	μ/κ	CLL #1	IGKV1-39*01	100	IGKJ2*02	CQQSYSTPPYTF
P14197	μ/κ	CLL #1	IGKV1-39*01	100.00	IGKJ4*01	CQQSYSTPTSF
P23728	μ/κ	CLL #1	IGKV1-39*01	100.00	IGKJ4*01	CQQSYSTPPLTF
P5220	μ/κ	CLL #1	IGKV1-39*01	100.00	IGKJ4*01	CQQSYSTPPLTF
P1173	μ/κ	CLL #1	IGKV1-39*01	100.00	GKJ2*01	CQQSYSTPPHTF
N4755	μ/λ	CLL #2	IGLV3-21*01	97.85	IGLJ3*02	CQVWDSGSDHPWVF
P11475	μ/λ	CLL #2	IGLV3-21*01	97.85	IGLJ3*02	CQVWDSGSDHPWVF
P326	μ/λ	CLL #2	IGLV3-21*01	99.7	IGLJ3*02	CQVWDSSSDHPWVF
CLL 412	μ/λ	CLL #2	IGLV3-21*01	99.28	IGLJ3*02	CQVWDSSSDHPWVF
KM1PL	μ/κ	CLL #6	IGKV3-20*01	100	IGKJ4*01	CQQYGSSPPLTF
P7150	μ/κ	CLL #6	IGKV3-20*01	100.00	IGKJ3*01	CQQYGSSPPFTF
P10824	μ/κ	CLL #6	IGKV3-20*01	100.00	IGKJ2*01	CQQYGSSPYTF
P11887	μ/κ	CLL #6	IGKV3-20*01	100.00	IGKJ1*01	CQQYGSSPPTF
P1430	μ/λ	CLL #7	IGLV3-9*01	99.64	IGLJ2*01	CQVWDSSTEKVF
P6275	μ/λ	CLL #7	IGLV3-10*01	99.64	IGLJ2*01	CYSTDSSGNPLF
P104	μ/λ	CLL #10	IGLV1-51*01	98.6	IGLJ1*01	CGTWDSSLSAYVF
P5071	μ/λ	CLL #10	IGLV1-40*01	98.96	IGLJ2*01	CQSYDSSLSVVF
BRNO_1520	μ/κ	CLL #13	IGKV3-20*01	95.74	IGKJ1*01	CQQYGNSPRTF
BRNO_0376	μ/κ	CLL #13	IGKV3-20*01	97.52	IGKJ2*02	CQQYGSSPRTF
BE064	μ/κ	CLL #59	IGKV2-28*01	100	IGKJ5*01	CMQALQTPITF
P5623	μ/λ	CLL #77	IGLV10-54*01	95.79	IGLJ2*01	CSAWDSSLSAQVF
P8407	μ/λ	CLL #77	IGLV10-54*01	95.79	IGLJ3*02	CSAWDSSLSARVF
P14438	μ/λ	CLL #77	IGLV10-54*01	95.79	IGLJ2*01	CSAWDNSLSAQIF
P2959	μ/λ	CLL #77	IGLV10-54*01	98.60	IGLJ3*02	CSAWDSSLSAWVF
P6360	μ/κ	CLL #148B	IGKV1-5*03	92.11	IGKJ5*01	CQQYTAFSVTF
P6540	μ/λ	CLL #169	IGLV3-21*01	98.21	IGLJ3*02	CQVWDSSSDHPWVF
P775	μ/λ	CLL #201	IGLV1-44*01	95.79	IGLJ3*02	CTAWDDSLSSWLF
L3961	μ/λ	CLL #201	IGLV1-44*01	96.49	IGLJ3*02	CAAWDDSLDAWLF
L7911	μ/λ	CLL #201	IGLV1-44*01	97.89	IGLJ3*02	CAAWDDSLNAWMF
P22229	μ/λ	CLL #201	IGLV1-44*01	95.44	IGLJ3*02	CAAWDDTLNAWLF
fungi_725	μ/κ	CLL #fungi[150]	IGKV2-24*01	97.96	IGKJ2*04	CMQATQICSL
fungi_3687	μ/κ	CLL #fungi[150]	IGKV2-24*01	97.96	IGKJ2*04	CMQATQLCSF

S.S.: CLL stereotyped subset

Result summary:	Productive IGH rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGHV1-2*04 F	score = 1435	identity = 100.00% (288/288 nt)
J-GENE and allele	Homsap IGHJ4*02 F	score = 240	identity = 100.00% (48/48 nt)
D-GENE and allele by IMGT/JunctionAnalysis	Homsap IGHD3-22*01 F	D-REGION is in reading frame 2	
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[25.17.38.11]	[8.8.16]	CARDQGGGSGYYHYFDYW

Result summary:	Productive IGK rearranged sequence: (no stop codon and in-frame junction)		
V-GENE and allele	Homsap IGKV1-39*01 F, or Homsap IGKV1D-39*01 F	score = 1390	identity = 100.00% (279/279 nt)
J-GENE and allele	Homsap IGKJ2*01 F	score = 154	identity = 89.47% (34/38 nt)
FR-IMGT lengths, CDR-IMGT lengths and AA JUNCTION	[26.17.36.10]	[6.3.10]	CQQSYSTPPDTF

***Figure 13**: Sequencing confirmation of the immunoglobulin genes insertion to the plasmid vectors. Analysis using IMGT/V-QUEST tool of the CLL plasmid carrying heavy (at the top) and light chain (at the bottom) rearrangements of patient P19994 immunoglobulin genes*

The co-transfection of HEK293T cells with IG heavy chain and corresponding IG light chain encoding plasmids resulted in highly efficient production of mAbs. In detail, individual transfections produced recombinant mAbs at concentrations ranging from 1,2 μg/ml to 20,8 μg/ml with an average concentration of 7 μg/ml. The integrity of the produced mAbs was evaluated by SDS gel electrophoresis and western blotting using anti-human IgM conjugated to HRP. (Figure 14)

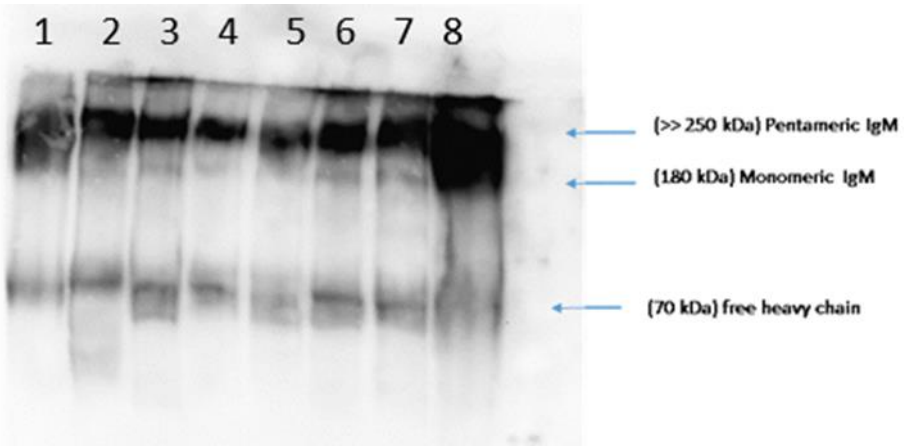


***Figure 14**: Representative example of SDS-PAGE of eight concentrated recombinant mAbs under non-reducing conditions. [(1:P411, 2:P8257, 3: P5623, 4:P8407, 5:P13836, 6:P22229, 7:L7911, 8:P23776) Western blotting was performed with goat anti-Human IgM HRP conjugated (abcam)]*

2. Antigen reactivity profiling

2.1 Reactivity against dsDNA

Seven SMZL and 29 CLL recombinant mAbs of the mu isotype were examined for dsDNA reactivity. Generally, the SMZL mAbs displayed stronger anti-dsDNA reactivity compared to CLL mAbs (Table 8, Figure 15). In detail, all 7 SMZL cases tested exhibited some binding ranging from low (2.5< OD ratio < 5, +) to very strong (OD ratio> 15, ++++). On the contrary, 21/29 (73%) CLL mAbs exhibited some anti-dsDNA reactivity that was characterized as either low (2.5< OD ratio < 5, +) or medium (5< OD ratio < 10, ++) for 17/29 (59%) and the 4/29 (14%) of the CLL mAbs tested, respectively.

Among the SMZL recombinant mAbs using the IGHV1-2*04, a wide range of dsDNA reactivity was observed. Cases P23776 and P13836, expressing minimally mutated or mutated IGHV1-2*04 genes, respectively, paired with mutated lambda light chains, exhibited stronger anti-dsDNA reactivity compared to: (i) case P411 carrying also a minimally mutated IGHV1-2*04 rearrangement, coupled however with a kappa light chain, and (ii) truly unmutated IGHV1-2*04 cases (cases P8761 and P19994, both paired with unmutated kappa light chains). Another case, namely P8257, expressing unmutated IGHV4-34 heavy chain, also exhibited strong binding to dsDNA (Figure 16).

Among CLL mAbs tested for anti-dsDNA reactivity in this study, only 2 cases belonging to stereotyped subset #1 (P14197 and P23728) and 2 cases belonging to stereotyped subset #201 (L7911 and P22229) exhibited binding of medium intensity (5< OD ratio < 10, ++). The remaining cases exhibited low (2.5< OD ratio < 5, +) if any (OD ratio < 2.5, -) reactivity (Figure 15). Importantly, cases belonging to the same stereotyped group exhibited similar profile of dsDNA binding (Table 8, Figures 15 and 17).

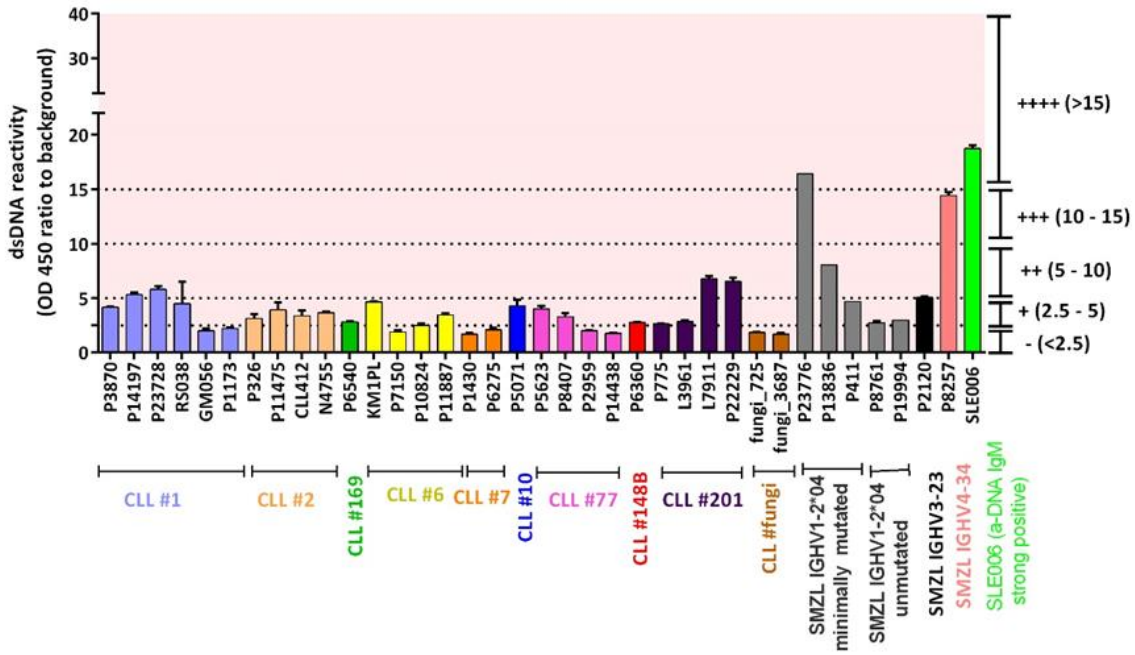


Figure 15: *Binding of CLL and SMZL IgM mAbs to dsDNA. The mAbs were used as primary mAbs in ELISA assays at a 10 µg/ml concentration. Serum of a patient with systemic lupus erythematosus (SLE 006) with anti-dsDNA IgM Abs, was used as positive control. Scale: signal-to-background ratio: – = < 2.5, + = 2.5-5, + + = 5-10, + + + = 10-15, + + + + = > 15.*

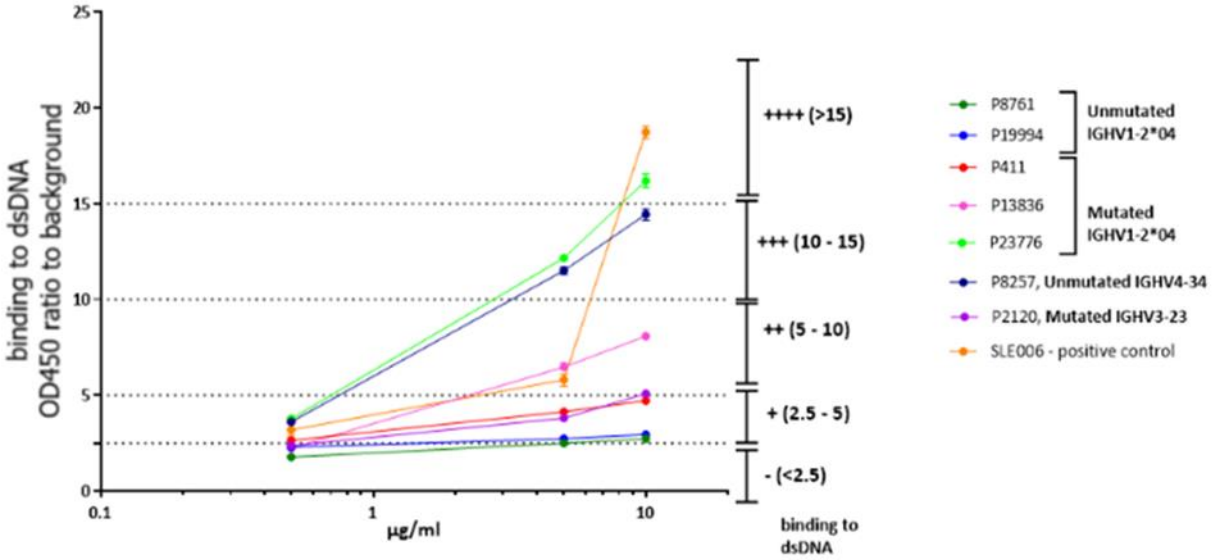


Figure 16: *Binding of SMZL IgM mAbs to dsDNA. The mAbs were used as primary mAbs in concentrations ranging from 0,5 µg/ml to 10 µg/ml. Serum of a patient with systemic lupus erythematosus (SLE 006) with anti-dsDNA IgM Abs, was used as positive control. Scale: signal-to-background ratio: – = < 2.5, + = 2.5-5, + + = 5-10, + + + = 10-15, + + + + = > 15.*

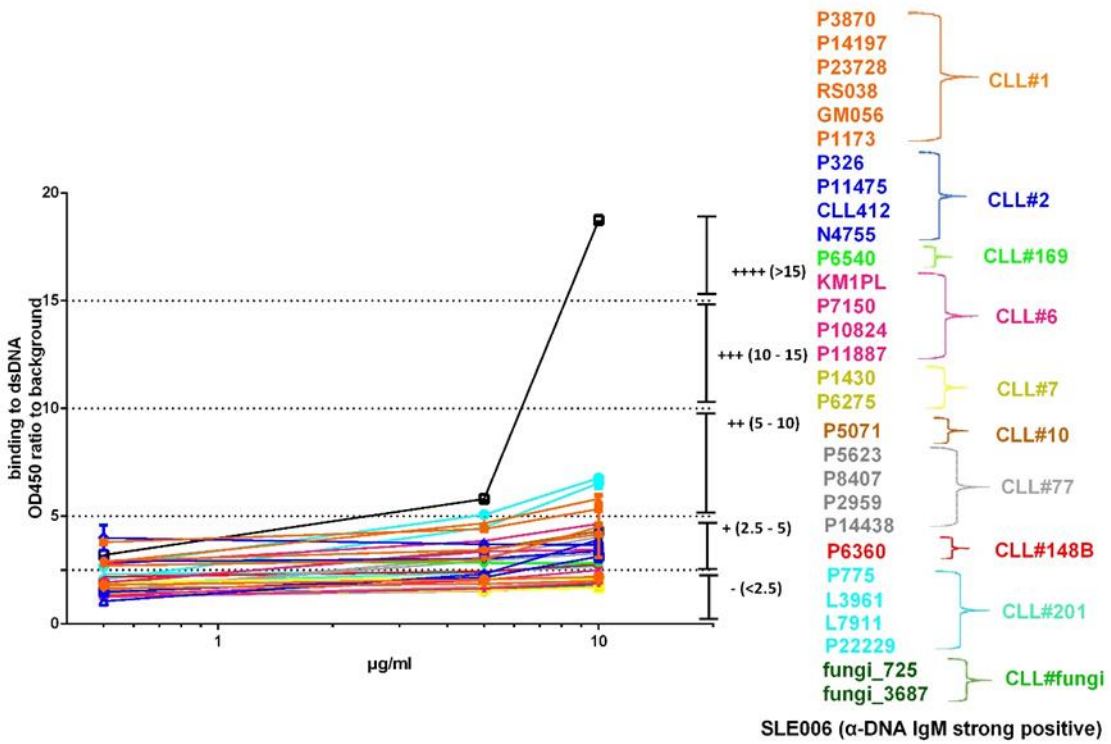


Figure 17: *Binding of CLL IgM mAbs to dsDNA. The mAbs were used as primary mAbs in concentrations ranging from 0,5 µg/ml to 10 µg/ml. Serum of a patient with systemic lupus erythematosus (SLE 006) with anti-dsDNA IgM Abs, was used as positive control. Scale: signal-to-background ratio: – = < 2.5, + = 2.5-5, + + = 5-10, + + + = 10-15, + + + + = > 15.*

2.2 ANA reactivity

ANA reactivity was generally variable with a tendency of higher reactivity in SMZL compared to CLL (Table 8, Figure 18). In detail, 6/7 tested SMZL cases exhibited some binding ranging from low (2.5< OD ratio < 5, +) to very strong (OD ratio> 15, ++++). On the contrary, CLL stereotyped subsets #77, #148B, #201, #fungi[150] were universally not reactive to ANA antigens (OD ratio < 2.5, -), while CLL stereotyped subsets #2 and #169 exhibited low ANA reactivity (2.5< OD ratio < 5, +). Cases belonging to CLL stereotyped subset #1 showed low (2.5< OD ratio < 5, +) or moderate (5< OD ratio < 10, ++) ANA reactivity.

Overall, the highest ANA reactivity was observed in CLL subsets #6, #7 and SMZL cases utilizing the IGHV1-2*04 and IGHV4-34 gene, albeit with high variability even amongst cases expressing the same gene (i.e. SMZL expressing IGHV1-2*04) or

belonging to the same stereotyped subset (CLL stereotyped subset #6 and #7) (Table 8, Figure 18).

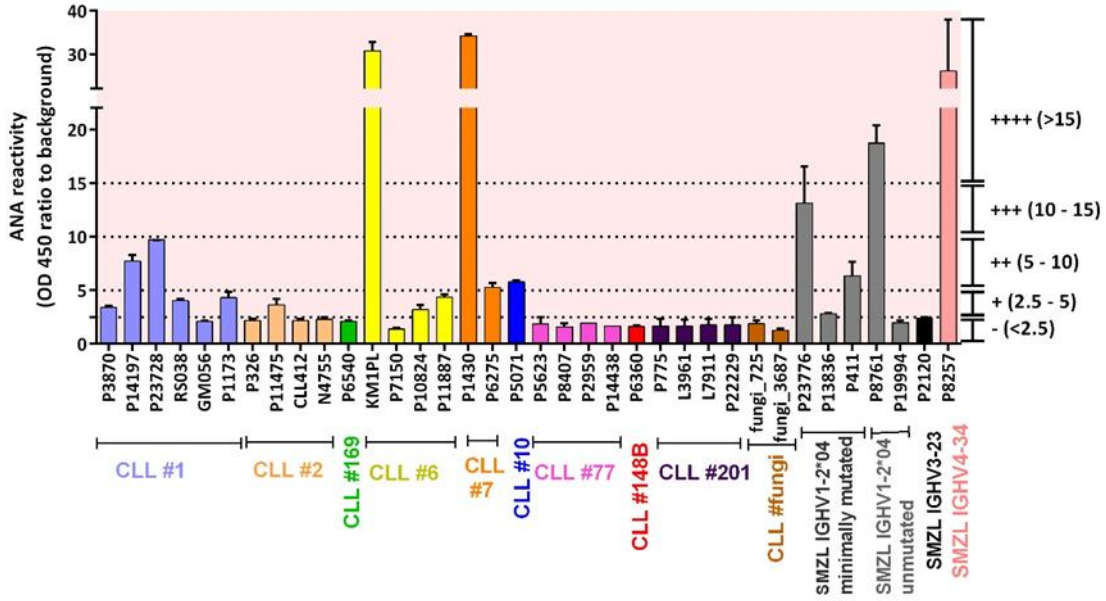


Figure 18: *Binding of CLL and SMZL IgM mAbs to nuclear antigens from Hep-2 cell extracts. Scale: signal-to-background ratio: – = < 2.5, + = 2.5-5, + + = 5-10, + + + = 10-15, + + + + = > 15.*

***Table 8**: ELISA assay results for the recognition of dsDNA and nuclear antigens of Hep-2 cell extracts by IgM mAbs. Scale: signal-to-background ratio: – = < 2.5, + = 2.5-5, + + = 5-10, + + + = 10-15, + + + + = > 15*

Case ID	Isotype	Entity and S.S.	IGHV gene	% gene Identity	dsDNA	nuclear antigens
P8761	μ/κ	SMZL	IGHV1-2*04	100.00	+	++++
P19994	μ/κ	SMZL	IGHV1-2*04	100.00	+	-
P411	μ/κ	SMZL	IGHV1-2*04	97.57	+	++
P23776	μ/λ	SMZL	IGHV1-2*04	98.96	++++	+++
P13836	μ/λ	SMZL	IGHV1-2*04	94.10	++	+
P2120	μ/κ	SMZL	IGHV3-23*01	97.92	++	+
P8257	μ/κ	SMZL	IGHV4-34*01	100.00	+++	++++
P3870	μ/κ	CLL #1	IGHV1-2*02	100.00	+	+
P14197	μ/κ	CLL #1	IGHV1-2*02	100.00	++	++
P23728	μ/κ	CLL #1	IGHV1-2*02	100.00	++	++
RS038	μ/κ	CLL #1	IGHV1-3*01	100.00	+	+
GM056	μ/κ	CLL #1	IGHV1-3*01	100.00	-	+
P1173	μ/κ	CLL #1	IGHV5-10-1*03	99.31	+	+
P326	μ/λ	CLL #2	IGHV3-21*02	97.92	+	+
P11475	μ/λ	CLL #2	IGHV3-21*01	97.99	+	+
CLL 412	μ/λ	CLL #2	IGHV3-21*01	98.25	+	+
N4755	μ/λ	CLL #2	IGHV3-21*01	97.92	+	+
P6540	μ/λ	CLL #169	IGHV3-48*02	96.18	+	+
KM1PL	μ/κ	CLL #6	IGHV1-69*06	100.00	+	++++
P7150	μ/κ	CLL #6	IGHV1-69*01	100.00	-	-
P10824	μ/κ	CLL #6	IGHV1-69*01	100.00	+	+
P11887	μ/κ	CLL #6	IGHV1-69*06	100.00	+	+
P1430	μ/λ	CLL #7	IGHV1-69*01	100.00	-	++++
P6275	μ/λ	CLL #7	IGHV1-69*01	100.00	-	++
P5071	μ/λ	CLL #10	IGHV4-39*01	100.00	+	++
P5623	μ/λ	CLL #77	IGHV4-59*03	91.93	+	-
P8407	μ/λ	CLL #77	IGHV4-4*02	92.01	+	-
P2959	μ/λ	CLL #77	IGHV4-4*02	95.09	-	-
P14438	μ/λ	CLL #77	IGHV4-59*01	89.47	-	-
P6360	μ/κ	CLL #148B	IGHV2-5*02	94.85	+	-
P775	μ/λ	CLL #201	IGHV4-34*01	94.74	+	-
L3961	μ/λ	CLL #201	IGHV4-34*01	91.23	+	-
L7911	μ/λ	CLL #201	IGHV4-34*01	94.39	++	-
P22229	μ/λ	CLL #201	IGHV4-34*01	91.58	++	-
fungi_725	μ/κ	CLL fungi[150]	IGHV3-7*01	95.49	-	-
fungi_3687	μ/κ	CLL fungi[150]	IGHV3-7*01	93.40	-	-

S.S.: CLL stereotyped subset according to Agathangelidis et al, 2012 and Hoogeboom et al, 2013

3. Binding of CLL and SMZL mAbs to viable MEC1 cells

Recent studies revealed that many CLL mAbs, mostly deriving from IG-unmutated CLL, recognize molecules on the surface of apoptotic cells, including human lymphocytes,[109,110]. Moreover, reactivity against viable cells has been detected in a IG-

mutated CLL subset.[151] With this in mind, we analyzed if CLL and SMZL clones can interact with viable lymphocytes, in particular with the MEC1 CLL cell line.[152]

To this end, we analyzed the binding of lambda-expressing mAbs only (Table 9). If Kappa-expressing mAbs were analyzed, a strong positive signal would not necessarily mean strong binding of recombinant mAbs to MEC1, since MEC1 cells express also surface IGs of the IgM/kappa isotype[152].

***Table 9**: mAbs examined for binding to viable MEC1 CLL cell line*

Case ID	Isotype	Entity/ S.S.	IGHV gene	% gene identity
P13836	μ/λ	SMZL	IGHV1-2*04	94.10
P23776	μ/λ	SMZL	IGHV1-2*04	98.96
L7911	μ/λ	CLL #201	IGHV4-34*01	94.39
L3961	μ/λ	CLL #201	IGHV4-34*01	91.23
P775	μ/λ	CLL #201	IGHV4-34*01	94.74
P22229	μ/λ	CLL #201	IGHV4-34*01	91.58
P5623	μ/λ	CLL #77	IGHV4-59*03	91.93
P8407	μ/λ	CLL #77	IGHV4-4*02	92.01
P2959	μ/λ	CLL #77	IGHV4-4*02	95.09
P14438	μ/λ	CLL #77	IGHV4-59*01	89.47
P326	μ/λ	CLL #2	IGHV3-21*02	97.92
P11475	μ/λ	CLL #2	IGHV3-21*01	97.99
CLL 412	μ/λ	CLL #2	IGHV3-21*01	98.25
N4755	μ/λ	CLL #2	IGHV3-21*01	97.92
P5071	μ/λ	CLL #10	IGHV4-39*01	100.00
P6540	μ/λ	CLL #169	IGHV3-48*02	96.18

S.S.: CLL stereotyped subset according to Agathangelidis et al, 2012

Only 7-AAD negative (viable cells) were included in the analysis (Figure 19).

The two SMZL cases utilizing the IGHV1-2*04 gene coupled with lambda light chain, were virtually non-reactive (Figure 20-21). In contrast, all CLL mAbs belonging to clinically indolent subset #201[153,154] bound the surfaces of MEC1 cells, with L7911 and L3961 being the most reactive. Interestingly these 2 cases but not the remaining 2 (P775 and P22229) carry novel N-glycosylation (N-glyc) motifs created by SHM within the VH FR3 region[153] (Figure 22A). Moreover, the differences between the #201 cases carrying or not novel N-glyc motifs at FR3 were statistically significant (Figure 22B).

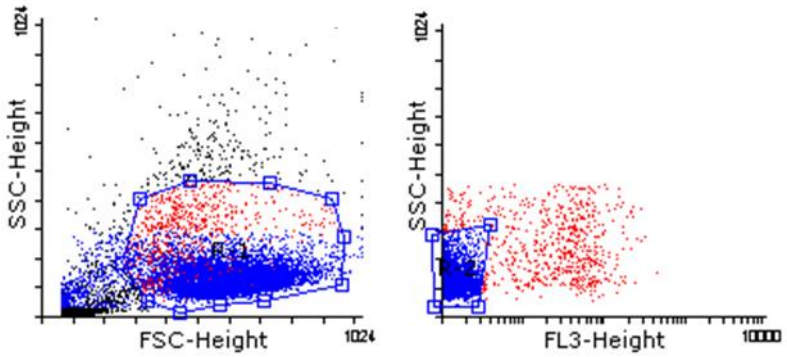


Figure 19: *Live MEC1 cell population was gated first from forward scatter and side scatter measurements (left panel) and then with the viability dye 7-AAD. FL3-Height represents the viability staining solution 7-AAD. (right panel, blue cells)*

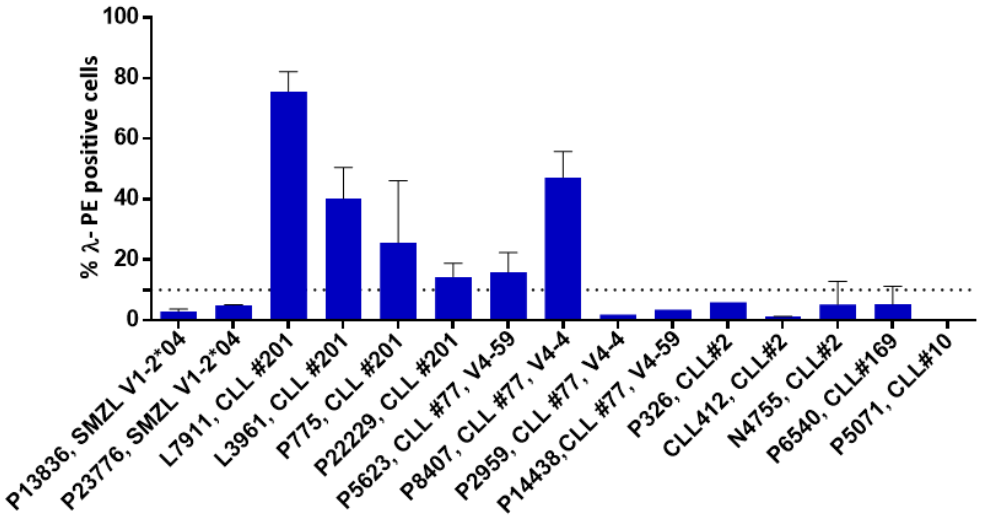


Figure 20: *Binding of SMZL and CLL IgM mAbs to viable MEC1 cells. Values are means from 2 or 3 separate experiments for each one of the cases tested. The error bars shown are standard deviations.*

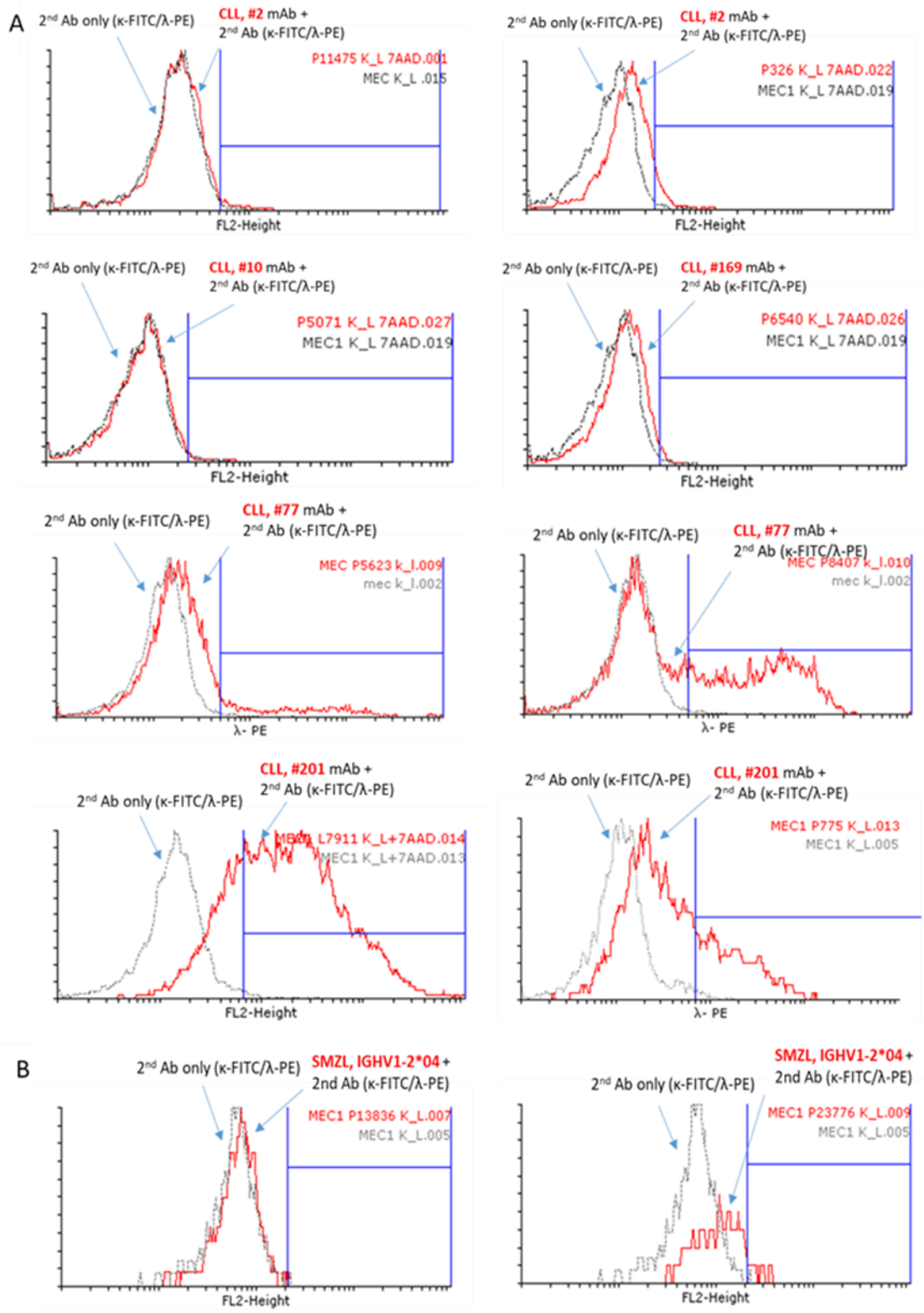


***Figure 21**: Examples of flow cytometry analysis for: A) eight CLL mAbs belonging to subsets #2, #10, #77, #169 or #201 and B) two SMZL mAbs utilizing the IGHV1-2*04 gene (P13836 and P23776). All mAbs have λ light chain and are expressed in IgM isotype. MEC1 cells with IgM κ-FITC/λ-PE were used as control.*

A

CLL #201 Case ID	N-glyc at FR3
L3961	YES
L7911	YES
P22229	NO
P775	NO

B

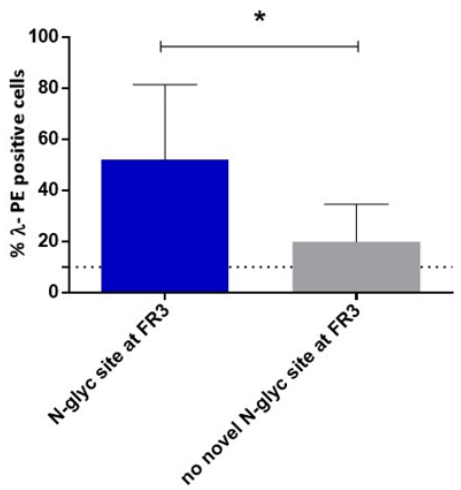


Figure 22: *A) Novel N-glycosylation (N-glyc) motifs created by SHM within the VH FR3 region of the clonotypic BcR IG of cases assigned to CLL stereotyped subset #201. B) CLL subset #201 cases with novel N-glyc motifs at the VH FR3 exhibit stronger binding to viable MEC1 cells. Values are means from three separate experiments for each one of the 4 cases tested (2 cases for each group). The error bars shown are standard deviations.*

Moreover, 2/4 cases (P8407, P5623) belonging to subset #77, also characterized by indolent clinical course, exhibited binding to viable MEC1.

At the opposite side of the clinical spectrum, aggressive stereotyped subset #2 and its immunogenetically closely related stereotyped subset #169[155] displayed no binding to viable MEC1 cell surfaces.

When CLL cases grouped on the basis of BcR stereotypy were compared, significant differences were observed. In particular, significantly higher interaction with viable MEC1 cells was observed in indolent CLL subsets #201 mAbs and #77, compared to subsets #2/#169. (Figure 23)

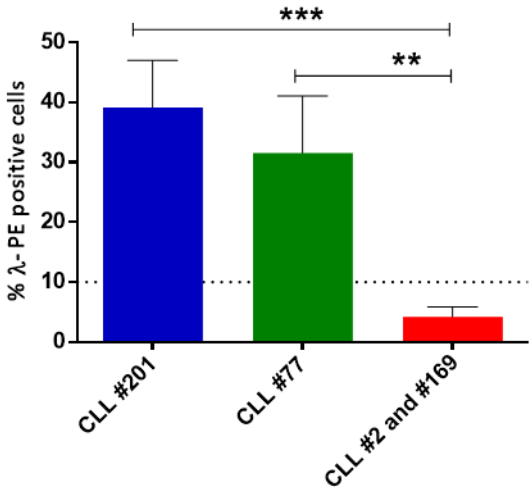


Figure 23: *Binding of CLL IgM/lambda mAbs to viable MEC1 cells. The studied cases were grouped on the basis of BCR stereotypy: n= 4, 4 and 5 for CLL #201, #77 and #2&169 respectively. Values are means from 2 or 3 separate experiments for each one of the cases tested. The error bars shown are standard deviations.*

4. Binding of CLL/SMZL mAbs to protein antigens from MEC1 cell extracts

To further characterize the binding to MEC1 cell epitopes, 4 different recombinant CLL mAbs belonging to subsets #4, #77, #201, all shown here or in previous studies[151] to interact with viable human B cells, as well as 2 SMZL cases were tested for recognition of membrane and cytoplasmic MEC1 protein antigens in a western blot analysis. (Table 10). The study group was enriched for cases encoded by the IGHV4-34 gene (4/6 cases): in particular, we analyzed the binding to MEC1 cell extracts of (i) 2 mutated, CLL subset #201 cases, (ii) 1 unmutated SMZL case and (iii) a mutated IgG IGHV4-34 mAb belonging to the indolent stereotyped subset #4[98] (case P907), produced previously in the lab in the context of another project. Non-IGHV4-34 expressing cases tested included an IGHV4-59 mAb belonging to subset #77 and a IGHV1-2*04 SMZL mAb.

Table 10*: Six mAbs tested for the binding to protein antigens from MEC1 cell extracts by Western Blot*

Case ID	Isotype	Entity/ CLL S.S	IGHV gene	% gene identity
L7911	μ/λ	CLL #201	IGHV4-34*01	94.39
P22229	μ/λ	CLL #201	IGHV4-34*01	91.58
P5623	μ/λ	CLL #77	IGHV4-59*03	91.93
P13836	μ/λ	SMZL	IGHV1-2*04	94.10
P8257	μ/κ	SMZL	IGHV4-34*01	100.00
P907*	γ/κ	CLL #4	IGHV4-34*02	93.20

** P907 case belongs to CLL stereotyped subset #4 defined by the expression of mutated IgHV4-33/IgKV2-30 BcR IGs of the γ isotype. The mAb was produced previously as a recombinant human IgG1.*

Blotting of MEC1 lysates with CLL/SMZL mAbs revealed recognition of protein antigens in 2/6 tested cases, both utilizing the IGHV4-34 gene. In particular, blotting with P8257 mAb, (SMZL unmutated IGHV4-34), revealed two bands at ~80 and ~220kDa, whereas P907 case, revealed two bands at ~220 and ~120kDa. (Figure 24)

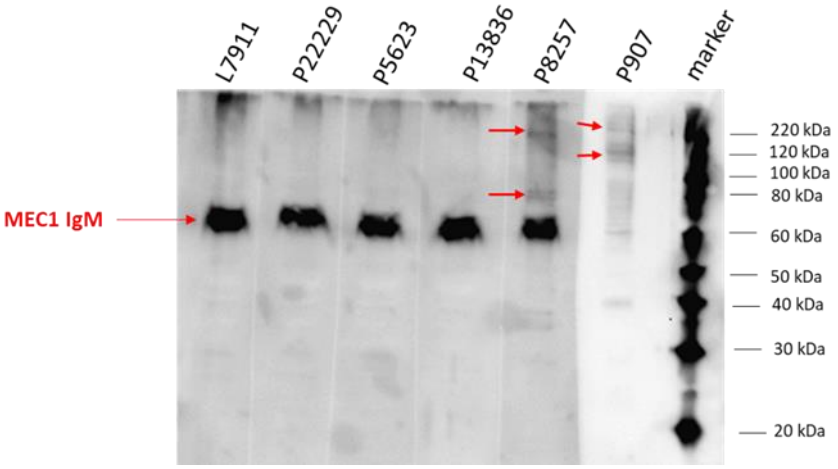


Figure 24: *Binding of CLL/SMZL mAbs to protein antigens from MEC1 cells in Western Blot analysis.*

Discussion

Accumulating evidence supports the notion that the BcR IG plays an important role in the ontogeny and evolution of malignancies of mature B cells (B cell non-Hodgkin lymphomas, B-NHLs).[156] From a clinical perspective, the importance of BcR IG-mediated signaling is highlighted by the remarkable clinical efficacy of B cell signaling inhibitors that interfere with the ability of B-NHL cells to sense microenvironmental stimuli.[157]

Immunogenetic support for microenvironmental interactions regarding CLL and SMZL derives from several observations, including immunogenetic findings. In particular: (i) the repertoire of clonotypic IGHV and IGK/LV genes is skewed[93,100], (ii) SHM is detected within the IG genes, often with distinctive imprints[95], (iii) intraclonal diversification due to ogoing somatic hypermutation is observed in certain subgroups of both CLL and SMZL attributed to continued pressure of antigens on the tumor cells[139] and (iv) unrelated patients express (quasi) identical BcR IG, a phenomenon called "BcR stereotypy"[100]. The latter is mainly observed in CLL, where 30% of CLL patients can be classified into a stereotyped subset.[100] Interestingly, CLL clones assigned to a distinct stereotyped subset are associated with distinct functional, biological but also clinical behavior.[102,158] This may be ascribed to the different ability of different CLL BcRs to bind their cognate antigens. On these grounds, it has been suggested that antigenic stimulation plays a crucial role in the ontogeny and evolution of B-NHL. Hence, the study of BcR IG antigenic reactivity has important implications for understanding the pathogenesis of these diseases. In particular, different processes initiated from the binding of the distinct leukemic BcR IG, that are often structurally biased, to their cognate antigen(s) could be involved in controlling the natural history of the disease, eventually determining the clinical outcome.

That notwithstanding, an alternative, non mutually exclusive scenario was put forth, according to which common immune-mediated mechanisms could underlie the emergence of distinct B-NHLs. Support to this concept was given by the identification of stereotyped IGHV4-59/IGKV3-20 receptors, identified in CLL and other lymphoproliferative conditions [SMZL, Monoclonal B cell lymphocytosis, HCV-associated type-II mixed cryoglobulinemia (HCV/MC-II), Sjögren's syndrome-

associated myoepithelial sialadenitis (SS-MESA)]. The stereotyped monoclonal IgM exhibited strong and specific rheumatoid factor activity, suggesting that specific microenvironment and/or similar if not identical antigenic elements select the clonogenic progenitors bearing this distinct BcR, resulting in distinct pathological conditions.[159]

Independent studies in CLL have revealed that the leukemic BcR IGs often exhibit polyspecificity and may recognize foreign antigens as well as self-antigens. In particular, autoantigens, molecular structures on microbial pathogens and neo-epitopes created by chemical modifications during apoptosis have been identified as antigenic targets for CLL. This indicates that CLL clones possibly arise from B cells which function as scavengers of apoptotic debris and also provide a first line of defense against pathogens, reminiscent of natural antibodies.[106–110] In addition, it has been demonstrated that unmutated CLL cases exhibit a higher degree of self-reactivity and polyreactivity compared to mutated CLL.[108,110] On the other hand, in SMZL, the single study published thus far concerned IGHV1-2*04 expressing clones and showed poly- and self-reactivity with recognition of nuclear, cytoplasmic and cell membrane antigens and human serum.[144] However, despite the progress achieved, the precise nature of the antigen(s) recognized by CLL, particularly M-CLL, and SMZL cases with different BcR IGs, is a still relatively unexplored area.[109–111,150,160]

In the present thesis, we examined the antigenic reactivity profile of (i) CLL BcR IGs originating from several stereotyped subsets of both unmutated and mutated CLL, associated with both favorable and adverse prognosis; and (ii) SMZL BcR IGs with a focus on cases expressing the most prevalent IGHV genes in the disease.[137]

The examination of the malignant BcR IGs presents important technical difficulties, because leukemic B cells have limited growth capacity *in vitro*, while CLL cells also express low levels of surface IG. In order to overcome these difficulties, various methods have been utilized such as stimulation of the malignant clone with PMA (phorbol myristate acetate), hybridoma technology or EBV transformation of leukemic cells.[106,107,161,162] However, EBV transformed B cells may be unstable and secrete small amounts of IGs.[163,164]

In 2008, Catera et al produced mAbs using recombinant DNA technology which is efficient regarding the IG production.[110] Nonetheless, all recombinant mAbs were

produced as human IgG1, thus they may differed from the actual leukemic IGs (that in most cases is of the mu isotype) regarding antigen affinity and avidity. In the present study, mAbs from 8 SMZL and 34 CLL patients were produced with recombinant DNA technology as human IgM, i.e. identical to the authentic heavy chain isotype, which is critical since the majority of leukemic clones of both CLL and SMZL express IGs of the mu isotype. Thus, it is more probable that the produced IgM mAbs have similar properties with the IGs expressed on the surface of the leukemic clone. Of note, several recent observations strongly suggest that the constant region can affect the specificity of the interactions between the Ab-binding site and the antigen.[165–167] Importantly, a recent structural study of the CLL BcR IG has shown that in certain cases the IG constant region is directly involved in the formation of the homotypic BcR interactions.[122]

Binding to nuclear antigens

Research on profiling the antigen reactivity of mAbs has demonstrated that an association between autoimmune phenomena and lymphoproliferative diseases may exist.[168–171] With this in mind, we initially examined the reactivity of recombinant mAbs from 7 SMZL (utilizing mutated or unmutated IGHV1-2*04, mutated IGHV3-23*01 and unmutated IGHV4-34*01) and 29 CLL cases (belonging to subsets: #1, #2, #6, #7, #10, #77, #148B, #201, #fungi)[98,150,153] against dsDNA and nuclear antigens from Hep-2 cell extracts, by ELISA assays. HEp-2 cells, originate from a human laryngeal carcinoma and allow recognition of nuclear and cytoplasmic patterns that are given by upwards of 50 different autoantibodies, correlated with various autoimmune conditions. The commercially available kit used in the present study, collectively detects total ANAs against dsDNA, histones, SS-A (Ro), SS-B (La), Sm, snRNP/Sm, Scl-70, PM-Scl, Jo-1 and centromeric antigens that occur in high frequency in systemic rheumatic diseases like systemic lupus erythematosus (SLE), mixed connective tissue diseases (MCTD), scleroderma, Sjögren`s syndrome, polymyositis and dermatomyositis.[172]

Samples from patients belonging to three CLL subsets, namely, #1, #2 and #7, as well as other minor CLL stereotyped subsets were included in the study. Subsets #1 and #2 express unmutated and borderline mutated BcR IGs[98,100] and are associated with aggressive disease.[154] As expected, mAbs from both subsets exhibited a low-medium

reactivity towards dsDNA and HEp-2 antigens. Interestingly, the reactivity profile of these IgM mAbs was similar with that of mAb from these subsets expressed as IgG1 and examined for binding to dsDNA and various nuclear antigens in a previous study.[110,160] Patients belonging to CLL subsets #6 and #7 usually present with aggressive disease and utilize the IGHV1-69 gene, which is the most common IGHV gene in CLL.[98] These two unmutated CLL subsets exhibited low, if any, dsDNA reactivity and variable ANA recognition, where 1 case from each subset revealed very strong ANA reactivity, contrasting the other cases within the same subset. On the contrary, subsets #77, #148B and #201 that have been associated with favorable prognosis[153,154], were universally not reactive to ANA antigens but certain cases showed low to medium dsDNA binding.

Clinically indolent subset #201[153] expresses BcR IG utilizing the IGHV4-34 gene. This gene encodes antibodies that are intrinsically autoreactive by virtue of universal, and largely light chain independent, recognition of the N-acetyllactosamine (NAL) antigenic determinant of the I/i blood group antigen; at least a subset of IGHV4-34 antibodies may also bind DNA.[173,174] IGHV4-34 antibodies are infrequent in the sera of healthy individuals, although the IGHV4-34 gene is very frequent in patients with systemic lupus erythematosus[118] and in the repertoire of peripheral B cells[173,174], suggesting an anergic status of these cells. Interestingly, IGHV4-34 is utilized also by the cases assigned to the stereotyped subset #4, the largest M-CLL subset and a prototype for indolent disease.[98,154] Unlike subset #201, that showed some anti-dsDNA reactivity, it has been demonstrated that cases belonging to subset #4, do not bind to dsDNA, likely due to censoring mechanism for avoiding intense self-reactivity, that introduced modifications by SHM.[151,175]

Regarding SMZL mAbs, stronger anti-dsDNA reactivity and a tendency for a higher ANA reactivity compared to CLL mAbs was observed. The high ANA reactivity that was observed in SMZL mAbs may be linked to the fact that approximately 20% of patients with SMZL present with autoimmune phenomena.[176] A special emphasis was given on cases expressing BcR IG utilizing the IGHV1-2*04 allele, the most commonly used allele in this lymphoma. Intriguingly, whereas almost all alleles of human IGHV genes carry an arginine (R; basic hydrophilic) at position VH FR3-75, IGHV1-2*04 carries a tryptophan (W; hydrophobic) at this exact position, which may alter the antibody conformation, thus having an impact on the recognition of specific antigenic

epitopes.[137] Among the SMZL recombinant mAbs that use the IGHV1-2*04, a wide range of dsDNA and ANA reactivity was observed. In a paper of 2011, Warsame et al suggested that SMZL using the IGHV1-02*04 mAbs are poly- and self-reactive recognizing nuclear, cytoplasmic and some membranous antigens expressed by human cells.[144] This is in accordance with our results, where we observed reactivity against nuclear antigens for SMZL mAbs expressing not only the IGHV1-2*04 allele but also other genes commonly expressed in SMZL (IGHV3-23 and IGHV4-34).[144]

Binding to B cell surfaces and protein extracts

Recent studies revealed that many CLL mAbs, mostly deriving from U-CLL, recognize molecules on the surface of apoptotic cells including human lymphocytes[109,110], albeit, reactivity also against viable cells has been detected in a M-CLL subset.[151] With this in mind, we analyzed if CLL and SMZL clones can interact with viable lymphocytes and in particular with MEC1 CLL cell line.[109,110,152] MEC1 cells express surface IGs of the IgM/kappa isotype; thus we analyzed the binding of lambda-expressing mAbs only (14 CLL mAbs belonging to subsets: #2, #10, #77, #169, #201 and 2 SMZL cases utilizing IGHV 1-2*04 paired with lambda light chain). If Kappa-expressing mAbs were analyzed, a strong positive signal would not necessarily mean strong binding of recombinant mAbs to MEC1, since MEC1 cells express also surface IGs of the IgM/kappa isotype.[152]

Amongst the 16 mAbs from lambda expressing cases tested for binding to the surface of MEC1 cells, reactivity against viable MEC1 cell surfaces was higher in subsets #201 and #77, both utilizing genes from the IGHV4 subgroup (IGHV4-34 and IGHV4-4/IGHV4-59, respectively). Interestingly binding to viable MEC1 cells in subsets #201 and #77, both associated with good prognosis,[153,154] was significantly higher compared to the clinically aggressive stereotyped subset #2.[154] Relevant to mention, mAbs from the extremely indolent stereotyped subset #4, also utilizing the IGHV4-34 gene, were recently shown to bind viable human B cell surfaces.[151]

Stereotyped subset #201 is also noteworthy owing to the high frequency of novel N-glycosylation motifs created by SHM within the VH FR3.[153] N-linked glycosylation is a widespread post-translational modification that can modulate antibody affinity towards antigen.[177,178] Interestingly, in #201 mAbs studied here, we observed a positive

correlation between the presence of novel N-glycosylation motifs within VH FR3 and binding to viable MEC1 cells.

SMZL mAbs produced in the present study did not recognize any structure on MEC1 cells. These results are in accordance with the study of Warsame et al, where the cell surface of lymphoma cell lines did not stain with the recombinant SMZL mAbs, utilizing the IGHV1-02 gene.[144]

To further characterize the binding to MEC1 cell epitopes, 4 different recombinant CLL mAbs belonging to subsets #4, #77, #201, all shown here or in previous studies[151] to interact with viable human B cells, as well as 1 SMZL case utilizing mutated IGHV1-2*04 and 1 unmutated IGHV4-34 expressing SMZL case, were also tested for recognition of membrane and cytoplasmic MEC1 protein antigens in a western blot analysis. Western blot analysis did not reveal any binding to MEC1 extracts for cases belonging to subsets #201 and #77 that bound to MEC1 cell surfaces as shown by flow cytometric analysis. This may allude to the recognition of either conformational protein antigens, that are eliminated by the denaturing conditions used in the immunoblotting experiments, by these mAbs. On the contrary, 2 other IGHV4-34 cases, an IgG switched CLL belonging to subset #4 and a SMZL mAb (P8257), recognized proteins expressed by MEC1 cells: probing with the mAb P8257 revealed 2 bands of ~80 and ~220kDa each, while probing with P907 mAb resulted in the identification of 2 predominant bands of ~120kDa and ~220kDa. The identification of a ~220kDa protein band after immunoblotting with these 2 mAbs both encoded by the IGHV4-34 gene implies that the main epitope of the band recognized could be shared by these 2 cases.

Final Conclusions

Altogether, our results further support the notion that grouping of CLL patients into subsets may be functionally and prognostically relevant, since cases belonging to the same subset exhibited similar antigen reactivity profile. Interestingly, while a low/medium anti-dsDNA reactivity was observed in many CLL cases, independently of the IGHV mutational status or the correlation with aggressive disease, ANA reactivity was observed only in CLL cases associated with an aggressive clinical course.

Moreover we showed that only the mAbs deriving from indolent CLL cases bound to viable MEC1 cells. In particular, mAbs from cases expressing the IGHV4-34 gene, that

is commonly used in both CLL and SMZL, bound strongly to viable lymphocytes, thus differing from the rest of CLL that usually recognize apoptotic cells.[109,110,145] The binding of viable lymphocytes by mAbs from the indolent CLL subset #201, as well as CLL subset #4 mAbs[151], both using the IGHV4-34 gene, suggests that the target antigen on the B cell surface is present constitutively and therefore differs from target antigens brought to the cell surface from the inside on apoptotic cells.[151]

It has been shown that subset #4 CLL B cells are anergic through the BcR[179] and this could be due to the chronic exposure to autoantigens[180,181] such as those expressed on the surface of viable lymphocytes. The indolent clinical course of subset #201 patients[153] allude to an anergic phenotype of the subset #201 CLL clone too, possibly mediated through similar procedures of chronic antigenic triggering. However, the differential imprints of SHM on IGHV4-34 stereotyped subsets, highlighted by the subset #201 biased introduction of N-glycosylation motifs by SHM, that we here show increases the affinity towards the antigen, allude to particular antigen exposure histories and/or immune responses.[98,153]

Further analysis is required in order to clarify the ontogenetic evolution of CLL BcR IGs, and in particular of those bearing SHM. This research should clarify the functional purpose for the changes introduced due to the SHM and unveil the antigens involved in the selection and/or evolution of B cell clones.[153]

Regarding SMZL mAbs, we generally observed stronger anti-dsDNA reactivity and a tendency for higher ANA reactivity compared to CLL mAbs. This observation implies differences in the antigen exposure history and/or in the subsequent immune responses to the (auto)antigen(s), of the malignant B cells and/or the pre-malignant B cells of these 2 distinct entities. The recognition of nuclear and cytoplasmic antigens by the SMZL mAbs analyzed, confirms and further extends the findings of a previous study, suggesting that SMZL B cell clones may arise from polyreactive B cells in the splenic marginal zone.[144]

Further analysis including more CLL and SMZL mAbs as well as a larger antigenic panel is required for a more precise characterization of the functional properties of the clonotypic BcR IGs. These analyses will lead to improved understanding of the role of BcR-antigen interactions in the natural history of B-NHLs.

BIBLIOGRAPHY

1. Janeway, C. A. & Medzhitov, R. Innate Immune Recognition. *Annu. Rev. Immunol.* **20,** 197–216 (2002).

2. Elves, P. J., Roitt, I. M., Mackay, I. R. & Rosen, F. S. The Immune System. *N. Engl. J. Med.* **343,** 37–49 (2000).

3. Panaroni, C., Tzeng, Y. S., Saeed, H. & Wu, J. Y. Mesenchymal progenitors and the osteoblast lineage in bone marrow hematopoietic niches. *Curr. Osteoporos. Rep.* **12,** 22–32 (2014).

4. Weissman, I. L. & Shizuru, J. A. The origins of the identification and isolation of hematopoietic stem cells, and their capability to induce donor-specific transplantation tolerance and treat autoimmune diseases. *Blood* **112,** 3543–3553 (2008).

5. Morrison, S. J. & Scadden, D. T. The bone marrow niche for haematopoietic stem cells. *Nature* **505,** 327–334 (2014).

6. Artis, D. & Spits, H. The biology of innate lymphoid cells. *Nature* **517,** 293–301 (2015).

7. Olszewski, W. L. The Lymphatic System in Body Homeostasis: Physiological Conditions. *Lymphat. Res. Biol.* **1,** 11–24 (2003).

8. Janeway CA Jr, Travers P, Walport M, et al. *Immunobiology: The Immune System in Health and Disease*. (New York: Garland Science, 2001).

9. Degn, S. E. & Thiel, S. Humoral Pattern Recognition and the Complement System. *Scand. J. Immunol.* **78,** 181–193 (2013).

10. Kawai, T. & Akira, S. The role of pattern-recognition receptors in innate immunity: Update on toll-like receptors. *Nat. Immunol.* **11,** 373–384 (2010).

11. Pasare, C. & Medzhitov, R. Toll-like receptors: linking innate and adaptive immunity. *Adv. Exp. Med. Biol.* **560,** 11–18 (2005).

12. Franchi, L., Warner, N., Viani, K. & Nuñez, G. Function of Nod-like Receptors in Microbial Recognition and Host Defense. *Immunol Rev.* **227,** 106–128 (2009).

13. Forthal, D. N. Functions of Antibodies. *Microbiol. Spectr.* **2,** 1–17 (2014).

14. Bonds, R. S., Midoro-Horiuti, T. & Goldblum, R. A structural basis for food allergy: the role of cross-reactivity. *Curr. Opin. Allergy Clin. Immunol.* **8,** 82–86 (2008).

15. Luft, F. C. Staphylococcus aureus, Toll-like receptors, superantigens, and their derivatives. *J. Mol. Med. (Berl).* **90,** 1091–1093 (2012).

16. Edelman, G. M. & Benacerraf, B. On structural and functional relations between antibodies and proteins of the gamma-system. *Proc. Natl. Acad. Sci. U. S. A.* **48,** 1035–1042 (1962).

17. Woof, J. M. & Burton, D. R. Human antibody–Fc receptor interactions illuminated by crystal structures. *Nat. Rev. Immunol.* **4,** 89–99 (2004).

18. Poljak, R. J. *et al.* X-Ray Crystallographic Studies of the Fab and Fc Fragments of Human Myeloma Immunoglobulins. *Cold Spring Harb Symp Quant Biol* **36,** 421–425 (1972).

19. Kabat, E. A. & Wu, T. T. Attempts To Locate Complementarity-Determining Residues in the Variable Positions of Light and Heavy Chains. *Ann. N. Y. Acad. Sci.* **190,** 382–393 (1971).

20. Lefranc, M. P. Nomenclature of the human immunoglobulin heavy (IGH) genes. *Exp. Clin. Immunogenet.* **18,** 100–116 (2001).

21. Lefranc, M. Variable (IGKV) Genes and Joining (IGKJ) Segments. *Exp Clin Immunogenet.* **5,** 171–183 (1998).

22. Lefranc, M. P. Nomenclature of the human immunoglobulin lambda (IGL) genes. *Exp. Clin. Immunogenet.* **18,** 242–254 (2001).

23. Schlissel, M. S. Regulating antigen-receptor gene assembly. *Nat. Rev. Immunol.* **3,** 890–899 (2003).

24. Gellert, M. V(D)J Recombination: RAG Proteins, Repair Factors, and Regulation. *Annu. Rev. Biochem.* **71,** 101–132 (2002).

25. Sadofsky, M. J. The RAG proteins in V(D)J recombination: more than just a nuclease. *Nucleic Acids Res.* **29,** 1399–1409 (2001).

26. Wei, Y. F. *et al.* Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase III, an enzyme active in DNA repair and recombination. *Mol. Cell. Biol.* **15,** 3206–3216 (1995).

27. Grawunder, U. *et al.* Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* **388,** 492–495 (1997).

28. Korsmeyer, S. J. *et al.* Developmental hierarchy of immunoglobulin gene rearrangements in human leukemic pre-B-cells. *Proc. Natl. Acad. Sci.* **78,** 7096–7100 (1981).

29. Vettermann, C. & Schlissel, M. S. Allelic exclusion of immunoglobulin genes: Models and mechanisms. *Immunol. Rev.* **237,** 22–42 (2010).

30. Papavasiliou, F. N. & Schatz, D. G. Somatic hypermutation of immunoglobulin genes: Merging mechanisms for genetic diversity. *Cell* **109,** 35–44 (2002).

31. Larson, E. D. & Maizels, N. Transcription-coupled mutagenesis by the DNA deaminase AID. *Genome Biol.* **5,** 211 (2004).

32. Muramatsu, M. *et al.* Specific Expression of Activation-induced Cytidine Deaminase (AID), a Novel Member of the RNA-editing Deaminase Family in Germinal Center B Cells. *J. Biol. Chem.* **274,** 18470–18476 (1999).

33. Bachl, J., Ertongur, I. & Jungnickel, B. Involvement of Rad18 in somatic hypermutation. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 12081–6 (2006).

34. Rogozin, I. B. & Pavlov, Y. I. Theoretical analysis of mutation hotspots and their

DNA sequence context specificity. *Mutat. Res. - Rev. Mutat. Res.* **544,** 65–85 (2003).

35. Shapiro, G. S. & Wysocki, L. J. DNA target motifs of somatic mutagenesis in antibody genes. *Crit. Rev. Immunol.* **22,** 183–200 (2002).

36. Smith, D. S. *et al.* Di- and trinucleotide target preferences of somatic mutagenesis in normal and autoreactive B cells. *J. Immunol.* **156,** 2642–2652 (1996).

37. Pan-Hammarström, Q., Zhao, Y. & Hammarström, L. Class Switch Recombination: A Comparison Between Mouse and Human. *Adv. Immunol.* **93,** 1–61 (2007).

38. Kenter, A. L. Class switch recombination: an emerging mechanism. *Immunity* **290,** 171–199 (2005).

39. Stavnezer-Nordgren, J. & Sirlin, S. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. *EMBO J.* **5,** 95–102 (1986).

40. Snapper, C. M., Marcu, K. B. & Zelazowski, P. The Immunoglobulin Class Switch : Beyond " Accessibility ". *Immunity* **6,** 217–223 (1997).

41. Iwasato, T., Shimizu, A. & Honjo, T. Circular DNA Is Excised by lmmunoglobulin Class Switch Recombination. *Cell* **62,** 143–149 (1990).

42. Muramatsu, M. *et al.* Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. *Cell* **102,** 553–563 (2000).

43. Chaudhuri, J. & Alt, F. W. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat. Rev. Immunol.* **4,** 655–655 (2004).

44. Kondo, M., Scherer, D. C., King, A. G., Manz, M. G. & Weissman, I. L. Lymphocyte development from hematopoietic stem cells. *Curr. Opin. Genet. Dev.* **11,** 520–526 (2001).

45. Hardy, R. R. & Hayakawa, K. B cell development pathways. *Annu Rev Immunol* **19,** 595–621 (2001).

46. Chowdhury, D. & Sen, R. Regulation of immunoglobulin heavy-chain gene rearrangements. *Immunol. Rev.* **200,** 182–196 (2004).

47. Monroe, J. G. & Dorshkind, K. Fate decisions regulating bone marrow and peripheral B lymphocyte development. *Adv. Immunol.* **95,** 1–50 (2007).

48. Alt, F. W., Blackwell, T. K. & Yancopoulos, G. D. Development of the primary antibody repertoire. *Science* **238,** 1079–1087 (1987).

49. Melchers, F. *et al.* The surrogate light chain in B-cell development. *Immunol. Today* **14,** 60–68 (1993).

50. Lassoued, K., Illges, H., Benlagha, K. & Cooper, M. D. Fate of surrogate light chains in B lineage cells. *J. Exp. Med.* **183,** 421–429 (1996).

51. Karasuyama, H., Rolink, A. & Melchers, F. Surrogate light chain in B cell development. *Adv. Immunol.* **63,** 1–41 (1996).

52. Keyna, U., Beck-Engeser, G. B., Jongstra, J., Applequist, S. E. & Jack, H. M. Surrogate light chain-dependent selection of Ig heavy chain V regions. *J. Immunol.* **155,** 5536–5542 (1995).

53. Kline, G. H. *et al.* Pre-B cell receptor-mediated selection of pre-B cells synthesizing functional mu heavy chains. *J. Immunol.* **161,** 1608–1618 (1998).

54. Melchers, F. Checkpoints that control B cell development. *J. Clin. Invest.* **125,** 2203–2210 (2015).

55. Ubelhart, R., Werner, M. & Jumaa, H. Assembly and Function of the Precursor B-Cell Receptor. *Curr. Top. Microbiol. Immunol.* **393,** 3–25 (2016).

56. Monroe, J. G. *et al.* Positive and negative selection during B lymphocyte development. *Immunol. Res.* **27,** 427–442 (2003).

57. Nemazee, D. Mechanisms of central tolerance for B cells. *Nat. Rev. Immunol.* **17,** 281–294 (2017).

58. Tiegs, S. L., Russell, D. M. & Nemazee, D. Receptor editing in self-reactive bone marrow B cells. *J. Exp. Med.* **177,** 1009–1020 (1993).

59. Radic, M. Z., Erikson, J., Litwin, S. & Weigert, M. B lymphocytes may escape tolerance by revising their antigen receptors. *J. Exp. Med.* **177,** 1165–1173 (1993).

60. Sims, G. P. *et al.* Identification and characterization of circulating human transitional B cells. *Blood* **105,** 4390–4398 (2005).

61. Pillai, S. & Cariappa, A. The follicular versus marginal zone B lymphocyte cell fate decision. *Nat. Rev. Immunol.* **9,** 767–777 (2009).

62. Cariappa, A. *et al.* Perisinusoidal B cells in the bone marrow participate in T-independent responses to blood-borne microbes. *Immunity* **23,** 397–407 (2005).

63. Cariappa, A., Chase, C., Liu, H., Russell, P. & Pillai, S. Naive recirculating B cells mature simultaneously in the spleen and bone marrow. *Blood* **109,** 2339–2345 (2007).

64. Gatto, D. & Brink, R. The germinal center reaction. *J. Allergy Clin. Immunol.* **126,** 898–899 (2010).

65. Pieper, K., Grimbacher, B. & Eibel, H. B-cell biology and development. *J. Allergy Clin. Immunol.* **131,** 959–971 (2013).

66. Pillai, S., Cariappa, A. & Moran, S. T. Marginal zone B cells. *Annu. Rev. Immunol.* **23,** 161–196 (2005).

67. Martin, F., Oliver, A. M. & Kearney, J. F. Marginal zone and B1 B cells unite in the early response against T-independent blood-borne particulate antigens. *Immunity* **14,** 617–629 (2001).

68. Cowan, M. J. *et al.* Pneumococcal polysaccharide immunization in infants and children. *Pediatrics* **62,** 721–727 (1978).

69. Likhite, V. V. Immunological impairment and susceptibility to infection after splenectomy. *JAMA* **236,** 1376–1377 (1976).

70. Attanavanich, K. & Kearney, J. F. Marginal zone, but not follicular B cells, are potent activators of naive CD4 T cells. *J. Immunol.* **172,** 803–811 (2004).

71. Berland, R. & Wortis, H. H. Origins and functions of B-1 cells with notes on the role of CD5. *Annu. Rev. Immunol.* **20,** 253–300 (2002).

72. Schwartz-Albiez, R., Monteiro, R. C., Rodriguez, M., Binder, C. J. & Shoenfeld, Y. Natural antibodies, intravenous immunoglobulin and their role in autoimmunity, cancer and inflammation. *Clin. Exp. Immunol.* **158,** 43–50 (2009).

73. Griffin, D. O., Holodick, N. E. & Rothstein, T. L. Human B1 cells in umbilical cord and adult peripheral blood express the novel phenotype CD20+ CD27+ CD43+ CD70-. *J. Exp. Med.* **208,** 67–80 (2011).

74. Rothstein, T. L., Griffin, D. O., Holodick, N. E., Quach, T. D. & Kaku, H. Human B-1 cells take the stage. *Ann. N. Y. Acad. Sci.* **1285,** 97–114 (2013).

75. Dal Porto, J. M. *et al.* B cell antigen receptor signaling 101. *Mol. Immunol.* **41,** 599–613 (2004).

76. Gold, M. R., Matsuuchi, L., Kelly, R. B. & DeFranco, A. L. Tyrosine phosphorylation of components of the B-cell antigen receptors following receptor crosslinking. *Proc. Natl. Acad. Sci. U. S. A.* **88,** 3436–3440 (1991).

77. Rowley, R. B., Burkhardt, A. L., Chao, H. G., Matsueda, G. R. & Bolen, J. B. Syk protein-tyrosine kinase is regulated by tyrosine-phosphorylated Ig alpha/Ig beta immunoreceptor tyrosine activation motif binding and autophosphorylation. *J. Biol. Chem.* **270,** 11590–11594 (1995).

78. Fruman, D. A. Phosphoinositide 3-kinase and its targets in B-cell and T-cell signaling. *Curr. Opin. Immunol.* **16,** 314–320 (2004).

79. Seda, V. & Mraz, M. B-cell receptor signalling and its crosstalk with other pathways in normal and malignant cells. *Eur. J. Haematol.* **94,** 193–205 (2015).

80. Satram-Hoang, S., Reyes, C., Hoang, K. Q., Momin, F. & Skettino, S. Treatment practice in the elderly patient with chronic lymphocytic leukemia-analysis of the combined SEER and Medicare database. *Ann. Hematol.* **93,** 1335–1344 (2014).

81. Matos, D. M. *et al.* Monoclonal B-cell lymphocytosis in first-degree relatives of patients with sporadic (non-familial) chronic lymphocytic leukaemia. *Br. J. Haematol.* **147,** 339–346 (2009).

82. Nabhan, C. *et al.* The impact of race, ethnicity, age and sex on clinical outcome in chronic lymphocytic leukemia: a comprehensive Surveillance, Epidemiology, and End Results analysis in the modern era. *Leuk. Lymphoma* **55,** 2778–2784 (2014).

83. Li, Y., Wang, Y., Wang, Z., Yi, D. & Ma, S. Racial differences in three major NHL subtypes: descriptive epidemiology. *Cancer Epidemiol.* **39,** 8–13 (2015).

84. Noone, A.-M. *et al.* Cancer Incidence and Survival Trends by Subtype Using

Data from the Surveillance Epidemiology and End Results Program, 1992-2013. *Cancer Epidemiol. Biomarkers Prev.* **26,** 632–641 (2017).

85. Redaelli, A., Laskin, B. L., Stephens, J. M., Botteman, M. F. & Pashos, C. L. The clinical and epidemiological burden of chronic lymphocytic leukaemia. *Eur. J. Cancer Care (Engl).* **13,** 279–287 (2004).

86. Hallek, M. *et al.* Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood* **111,** 5446–5456 (2008).

87. Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, P. B. Clinical staging of chronic lymphocytic leukemia. *Blood* **46,** 219–234 (1975).

88. Binet, J. L. *et al.* A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. *Cancer* **48,** 198–206 (1981).

89. Dicker, F., Schnittger, S., Haferlach, T., Kern, W. & Schoch, C. Immunostimulatory oligonucleotide-induced metaphase cytogenetics detect chromosomal aberrations in 80% of CLL patients: A study of 132 CLL cases with correlation to FISH, IgVH status, and CD38 expression. *Blood* **108,** 3152–3160 (2006).

90. Puiggros, A., Blanco, G. & Espinet, B. Genetic abnormalities in chronic lymphocytic leukemia: where we are and where we go. *Biomed Res. Int.* **2014,** 435983 (2014).

91. Puente, X. S. *et al.* Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **475,** 101–105 (2011).

92. Gruber, M. & Wu, C. J. Evolving understanding of the CLL genome. *Semin. Hematol.* **51,** 177–187 (2014).

93. Fais, F. *et al.* Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. *J. Clin. Invest.* **102,** 1515–1525 (1998).

94. Hamblin, T. J., Davis, Z., Gardiner, A., Oscier, D. G. & Stevenson, F. K. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* **94,** 1848–54 (1999).

95. Damle, R. N. *et al.* Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* **94,** 1840–1847 (1999).

96. McHeyzer-Williams, M., Okitsu, S., Wang, N. & McHeyzer-Williams, L. Molecular programming of B cell memory. *Nat. Rev. Immunol.* **12,** 24–34 (2011).

97. Potter, K. N. *et al.* Features of the overexpressed V1-69 genes in the unmutated subset of chronic lymphocytic leukemia are distinct from those in the healthy elderly repertoire. *Blood* **101,** 3082–3084 (2003).

98. Stamatopoulos, K. & Belessi, C. Over 20% of patients with chronic lymphocytic

leukemia carry stereotyped receptors: pathogenetic implications and clinical correlations. *Blood* **109,** 259–270 (2007).

99. Messmer, B. T. *et al.* Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *J. Exp. Med.* **200,** 519–525 (2004).

100. Agathangelidis, A. *et al.* Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: A molecular classification with implications for targeted therapies. *Blood* **119,** 4467–4475 (2012).

101. Murray, F. *et al.* Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood* **111,** 1524–1533 (2008).

102. Stamatopoulos, K., Agathangelidis, A., Rosenquist, R. & Ghia, P. Antigen receptor stereotypy in chronic lymphocytic leukemia. *Leukemia* **31,** 282–291 (2017).

103. Stevenson, F. K., Forconi, F. & Packham, G. The meaning and relevance of B-cell receptor structure and function in chronic lymphocytic leukemia. *Semin. Hematol.* **51,** 158–167 (2014).

104. Klein, U. *et al.* Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J. Exp. Med.* **194,** 1625–1638 (2001).

105. Mockridge, C. I. *et al.* Reversible anergy of sIgM-mediated signaling in the two subsets of CLL defined by VH-gene mutational status. *Blood* **109,** 4424–4431 (2007).

106. Sthoeger, Z. M. *et al.* Production of autoantibodies by CD5-expressing B lymphocytes from patients with chronic lymphocytic leukemia. *J. Exp. Med.* **169,** 255–268 (1989).

107. Borche, L., Lim, A., Binet, J. L. & Dighiero, G. Evidence that chronic lymphocytic leukemia B lymphocytes are frequently committed to production of natural autoantibodies. *Blood* **76,** 562–569 (1990).

108. Herve, M. *et al.* Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. *J. Clin. Invest.* **115,** 1636–1643 (2005).

109. Lanemo Myhrinder, A. *et al.* A new perspective: molecular motifs on oxidized LDL, apoptotic cells, and bacteria are targets for chronic lymphocytic leukemia antibodies. *Blood* **111,** 3838–3848 (2008).

110. Catera, R. *et al.* Chronic lymphocytic leukemia cells recognize conserved epitopes associated with apoptosis and oxidation. *Mol. Med.* **14,** 665–674 (2008).

111. Chu, C. C. *et al.* Chronic lymphocytic leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. *Blood* **112,** 5122–5129 (2008).

112. Chang, M.-K. *et al.* Apoptotic cells with oxidation-specific epitopes are immunogenic and proinflammatory. *J. Exp. Med.* **200,** 1359–1370 (2004).

113. Hacken, E. ten *et al.* Calreticulin as a novel B-cell receptor antigen in chronic lymphocytic leukemia. *Haematologica* **102,** e394–e396 (2017).

114. Hatzi, K. *et al.* B-Cell Chronic Lymphocytic Leukemia (B-CLL) Cells Express Antibodies Reactive with Antigenic Epitopes Expressed on the Surface of Common Bacteria. *Blood* **108,** 25 LP-25 (2006).

115. Landgren, O. *et al.* Respiratory tract infections and subsequent risk of chronic lymphocytic leukemia. *Blood* **109,** 2198–2201 (2007).

116. Hatzi, K. *et al.* Chronic lymphocytic leukemia immunoglobulins display bacterial reactivity that converges and diverges from auto-/poly-reactivity and IGHV mutation status. *Clin. Immunol.* **172,** 44–51 (2016).

117. Kostareli, E. *et al.* Molecular evidence for EBV and CMV persistence in a subset of patients with chronic lymphocytic leukemia expressing stereotyped IGHV4-34 B-cell receptors. *Leukemia* **23,** 919–924 (2009).

118. Mockridge, C. I. *et al.* Common patterns of B cell perturbation and expanded V4-34 immunoglobulin gene usage in autoimmunity and infection. *Autoimmunity* **37,** 9–15 (2004).

119. Bhat, N. M. *et al.* B cell lymphoproliferative disorders and VH4-34 gene encoded antibodies. *Hum. Antibodies* **13,** 63–68 (2004).

120. Li, M. *et al.* Hepatitis B virus and risk of non-Hodgkin lymphoma: An updated meta-analysis of 58 studies. *J. Viral Hepat.* (2018). doi:10.1111/jvh.12892

121. Keyvani, H., Karbalaie Niya, M. H., Esghaei, M., Bokharaei-Salim, F. & Monavari, S. H. R. Presence of human herpesvirus 8 (HHV-8) DNA sequences in patients with lymphoproliferative diseases and chronic blood disorders. *Microb. Pathog.* **111,** 431–434 (2017).

122. Minici, C. *et al.* Distinct homotypic B-cell receptor interactions shape the outcome of chronic lymphocytic leukaemia. *Nat. Commun.* **8,** 15746 (2017).

123. Duhren-von Minden, M. *et al.* Chronic lymphocytic leukaemia is driven by antigen-independent cell-autonomous signalling. *Nature* **489,** 309–312 (2012).

124. World Health Organization. Classification of Tumours. Pathology and genetics of tumours of haematopoietic and lymphoid tissues. IARC Press: Lyon, 2001. doi:10.1016/j.hoc.2008.07.011

125. Olszewski, A. J. & Castillo, J. J. Survival of patients with marginal zone lymphoma: analysis of the Surveillance, Epidemiology, and End Results database. *Cancer* **119,** 629–638 (2013).

126. Arcaini, L., Rossi, D. & Paulli, M. Splenic marginal zone lymphoma: from genetics to management. *Blood* **127,** 2072–2081 (2016).

127. Liu, L. *et al.* Splenic marginal zone lymphoma: a population-based study on the 2001-2008 incidence and survival in the United States. *Leuk. Lymphoma* **54,** 1380–1386 (2013).

128. de Sanjose, S. *et al.* Hepatitis C and non-Hodgkin lymphoma among 4784 cases and 6269 controls from the International Lymphoma Epidemiology Consortium.

Clin. Gastroenterol. Hepatol. **6,** 451–458 (2008).

129. Thieblemont, C. *et al.* Treatment of splenic marginal zone B-cell lymphoma: an analysis of 81 patients. *Clin. Lymphoma* **3,** 41–47 (2002).

130. Matutes, E. *et al.* Splenic marginal zone lymphoma proposals for a revision of diagnostic, staging and therapeutic criteria. *Leukemia* **22,** 487–495 (2008).

131. Baliakas, P. *et al.* Splenic marginal-zone lymphoma: Ontogeny and genetics. *Leuk. Lymphoma* **56,** 301–310 (2015).

132. Salido, M. *et al.* Cytogenetic aberrations and their prognostic value in a series of 330 splenic marginal zone B-cell lymphomas: a multicenter study of the Splenic B-Cell Lymphoma Group. *Blood* **116,** 1479–1488 (2010).

133. Kiel, M. J. *et al.* Whole-genome sequencing identifies recurrent somatic NOTCH2 mutations in splenic marginal zone lymphoma. *J. Exp. Med.* **209,** 1553–1565 (2012).

134. Parry, M. *et al.* Genetics and Prognostication in Splenic Marginal Zone Lymphoma: Revelations from Deep Sequencing. *Clin. Cancer Res.* **21,** 4174–4183 (2015).

135. Spina, V. & Rossi, D. Molecular pathogenesis of splenic and nodal marginal zone lymphoma. *Best Pract. Res. Clin. Haematol.* **30,** 5–12 (2017).

136. Rossi, D. *et al.* The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J. Exp. Med.* **209,** 1537–1551 (2012).

137. Bikos, V. *et al.* Over 30% of patients with splenic marginal zone lymphoma express the same immunoglobulin heavy variable gene: Ontogenetic implications. *Leukemia* **26,** 1638–1646 (2012).

138. Bikos, V. *et al.* Selection of antigen receptors in splenic marginal-zone lymphoma: Further support from the analysis of the immunoglobulin light-chain gene repertoire. *Leukemia* **26,** 2567–2569 (2012).

139. Bikos, V. *et al.* An immunogenetic signature of ongoing antigen interactions in splenic marginal zone lymphoma expressing IGHV1-2∗04 receptors. *Clin. Cancer Res.* **22,** 2032–2040 (2016).

140. Arcaini, L. *et al.* Splenic and nodal marginal zone lymphomas are indolent disorders at high hepatitis C virus seroprevalence with distinct presenting features but similar morphologic and phenotypic profiles. *Cancer* **100,** 107–115 (2004).

141. Vallisa, D. *et al.* Role of anti-hepatitis C virus (HCV) treatment in HCV-related, low-grade, B-cell, non-Hodgkin's lymphoma: a multicenter Italian experience. *J. Clin. Oncol.* **23,** 468–473 (2005).

142. Mele, A. *et al.* Hepatitis C virus and B-cell non-Hodgkin lymphomas: an Italian multicenter case-control study. *Blood* **102,** 996–999 (2003).

143. Bates, I., Bedu-Addo, G., Rutherford, T. & Bevan, D. H. Splenic lymphoma with villous lymphocytes in tropical West Africa. *Lancet (London, England)* **340,**

575–577 (1992).

144. Warsame, A. A. *et al.* Splenic marginal zone lymphoma with VH1-02 gene rearrangement expresses poly- and self-reactive antibodies with similar reactivity. *Blood* **118,** 3331–3339 (2011).

145. Chu, C. C. *et al.* Many chronic lymphocytic leukemia antibodies recognize apoptotic cells with exposed nonmuscle myosin heavy chain IIA: implications for patient outcome and cell of origin. *Blood* **115,** 3907–3915 (2010).

146. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular cloning: a laboratory manual.* (Cold Spring Harbor Laboratory Press, 1989).

147. Dighiero, G. *et al.* B-cell chronic lymphocytic leukemia: present status and future directions. French Cooperative Group on CLL. *Blood* **78,** 1901 LP-1914 (1991).

148. Arcaini, L. & Paulli, M. Splenic marginal zone lymphoma: hydra with many heads? *Haematologica* **95,** 534–537 (2010).

149. Tiller, T. *et al.* Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J. Immunol. Methods* **329,** 112–124 (2008).

150. Hoogeboom, R. *et al.* A mutated B cell chronic lymphocytic leukemia subset that recognizes and responds to fungi. *J. Exp. Med.* **210,** 59–70 (2013).

151. Catera, R. *et al.* Binding of CLL subset 4 B-cell receptor immunoglobulins to viable human memory B lymphocytes requires a distinctive IGKV somatic mutation. *Mol. Med.* **23,** (2017).

152. Stacchini, A. *et al.* MEC1 and MEC2: two new cell lines derived from B-chronic lymphocytic leukaemia in prolymphocytoid transformation. *Leuk. Res.* **23,** 127–136 (1999).

153. Xochelli, A. *et al.* Chronic lymphocytic leukemia with mutated IGHV4-34 receptors: Shared and distinct immunogenetic features and clinical outcomes. *Clin. Cancer Res.* **23,** 5292–5301 (2017).

154. Baliakas, P. *et al.* Clinical effect of stereotyped B-cell receptor immunoglobulins in chronic lymphocytic leukaemia: a retrospective multicentre study. *Lancet. Haematol.* **1,** e74-84 (2014).

155. Agathangelidis, A. *et al. Higher-Order Immunoglobulin sequence relations for major subsets of Chronic Lymphocytic Leukemia: uniqueness versus equivalence*. *Haematologica* **100,** (2015).

156. Suarez, F. & Lecuit, M. Infection-associated non-Hodgkin lymphomas. *Clin. Microbiol. Infect.* **21,** 991–997 (2015).

157. Wiestner, A. Targeting B-Cell receptor signaling for anticancer therapy: the Bruton's tyrosine kinase inhibitor ibrutinib induces impressive responses in B-cell malignancies. *J. Clin. Oncol.* **31,** 128–130 (2013).

158. Rossi, D. & Gaidano, G. Biological and clinical significance of stereotyped B-cell receptors in chronic lymphocytic leukemia. *Haematologica* **95,** 1992–1995

(2010).

159. Kostareli, E. *et al.* Antigen receptor stereotypy across B-cell lymphoproliferations: the case of IGHV4-59/IGKV3-20 receptors with rheumatoid factor activity. *Leukemia* **26,** 1127–1131 (2012).

160. Gounari, M. *et al.* Excessive antigen reactivity may underlie the clinical aggressiveness of chronic lymphocytic leukemia stereotyped subset #8. *Blood* **125,** 3580–3587 (2015).

161. Broker, B. M. *et al.* Chronic lymphocytic leukemic (CLL) cells secrete multispecific autoantibodies. *J. Autoimmun.* **1,** 469–481 (1988).

162. Jahn, S. *et al.* Human hybridomas derived from CD5+ B lymphocytes of patients with chronic lymphocytic leukemia (B-CLL) produce multi-specific natural IgM (kappa) antibodies. *Clin. Exp. Immunol.* **83,** 413–417 (1991).

163. Siemer, D. *et al.* EBV transformation overrides gene expression patterns of B cell differentiation stages. *Mol. Immunol.* **45,** 3133–3141 (2008).

164. James, K. & Bell, G. T. Human monoclonal antibody production. Current status and future prospects. *J. Immunol. Methods* **100,** 5–40 (1987).

165. Feng, J. Q., Mozdzanowska, K. & Gerhard, W. Complement component C1q enhances the biological activity of influenza virus hemagglutinin-specific antibodies depending on their fine antigen specificity and heavy-chain isotype. *J. Virol.* **76,** 1369–1378 (2002).

166. Cooper, L. J. *et al.* Role of heavy chain constant domains in antibody-antigen interaction. Apparent specificity differences among streptococcal IgG antibodies expressing identical variable domains. *J. Immunol.* **150,** 2231–2242 (1993).

167. Torres, M., May, R., Scharff, M. D. & Casadevall, A. Variable-region-identical antibodies differing in isotype demonstrate differences in fine specificity and idiotype. *J. Immunol.* **174,** 2132–2142 (2005).

168. Duek, A., Shvidel, L., Braester, A. & Berrebi, A. Clinical and immunologic aspects of B chronic lymphocytic leukemia associated with autoimmune disorders. *Isr. Med. Assoc. J.* **8,** 828–831 (2006).

169. Barcellini, W. *et al.* Relationship between autoimmune phenomena and disease stage and therapy in B-cell chronic lymphocytic leukemia. *Haematologica* **91,** 1689–1692 (2006).

170. Vanura, K. *et al.* Autoimmune conditions and chronic infections in chronic lymphocytic leukemia patients at diagnosis are associated with unmutated IgVH genes. *Haematologica* **93,** 1912–1916 (2008).

171. Voinchet, H. *et al.* [Splenic marginal zone lymphoma and autoimmunity: report of six cases]. *La Rev. Med. interne* **31,** 4–11 (2010).

172. Meroni, P. L. & Schur, P. H. ANA screening: an old test with new recommendations. *Ann. Rheum. Dis.* **69,** 1420–1422 (2010).

173. Pugh-Bernard, A. E. *et al.* Regulation of inherently autoreactive VH4-34 B cells in the maintenance of human B cell tolerance. *J. Clin. Invest.* **108,** 1061–1070

(2001).

174. Silberstein, L. E., George, A., Durdik, J. M. & Kipps, T. J. The V4-34 encoded anti-i autoantibodies recognize a large subset of human and mouse B-cells. *Blood Cells. Mol. Dis.* **22,** 126–138 (1996).

175. Catera, R. *et al.* Somatic Hypermutation In Stereotyped Subset 4 BCRs/mAbs of CLL Patients, Expressing IGHV4-34 gene, Edit Anti-DNA Reactivity. *Blood* **116,** 2444 LP-2444 (2010).

176. Santos, T. S. Dos, Tavares, R. S. & Farias, D. L. C. de. Splenic marginal zone lymphoma: a literature review of diagnostic and therapeutic challenges. *Rev. Bras. Hematol. Hemoter.* **39,** 146–154 (2017).

177. Garces, F. *et al.* Affinity maturation of a potent family of HIV antibodies is primarily focused on accommodating or avoiding glycans. *Immunity* **43,** 1053–1063 (2015).

178. Wallick, S. C., Kabat, E. A. & Morrison, S. L. Glycosylation of a VH residue of a monoclonal antibody against alpha (1----6) dextran increases its affinity for antigen. *J. Exp. Med.* **168,** 1099 LP-1109 (1988).

179. Ntoufa, S. *et al.* B Cell Anergy Modulated by TLR1/2 and the miR-17 approximately 92 Cluster Underlies the Indolent Clinical Course of Chronic Lymphocytic Leukemia Stereotyped Subset #4. *J. Immunol.* **196,** 4410–4417 (2016).

180. Healy, J. I. *et al.* Different nuclear signals are activated by the B cell receptor during positive versus negative signaling. *Immunity* **6,** 419–428 (1997).

181. Cambier, J. C., Gauld, S. B., Merrell, K. T. & Vilen, B. J. B-cell anergy: from transgenic models to naturally occurring anergic B cells? *Nat. Rev. Immunol.* **7,** 633–643 (2007).