ANALYSIS OF ENDOGENOUS IMMUNOGLOBULIN GENE REARRANGEMENT IN THE SCID MOUSE

by

Jacqueline Louise Mitchell Hamilton Pennycook

A thesis submitted in conformity with the requirements for the Degree of Doctor of Philosophy, Graduate Department of Immunology, in the University of Toronto.

© by Jacqueline L. M. H. Pennycook, 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-27711-9
ABSTRACT

Analysis of Endogenous Immunoglobulin Gene Rearrangement in the SCID Mouse.

Degree of Doctor of Philosophy, 1997.

Jacqueline Louise Mitchell Hamilton Pennycook

Graduate Department of Immunology, University of Toronto

The critical event in lymphoid development is the assembly of germline variable gene segments forming the variable exon of the antigen receptors. This somatic assembly occurs through a process called V(D)J recombination. Mice with severe combined immune deficiency (SCID) have a profound deficiency in lymphoid development. The scid mutation results in a defect in the ability to repair double strand DNA damage, and consequently affects normal V(D)J rearrangement. Early studies showed that endogenous rearrangements in transformed SCID cell lines exhibited large deletions in the recombined genetic elements. Indeed, the estimated frequency of normal coding joint resolution in SCID cells is $10^2$- to $10^3$-fold lower than wild type. However, the scid mutation in not absolute as older mice become "leaky", exhibiting serum immunoglobulin (Ig) and/or T cell activity. During the course of my studies, the SCID gene product was identified as the catalytic subunit of the DNA-dependent protein kinase (DNA-PKcs); a molecule implicated in ds DNA damage repair, V(D)J recombination and transcription regulation.

The goal of my thesis is to acquire greater understanding of the process of V(D)J recombination and its role in B cell development. To achieve this objective, I have examined the extent of assembly of Ig variable genes in non-
leaky SCID mice. In Chapter III, I show that DJH rearrangements are found at 1-10% the frequency observed in C.B-17 whole bone marrow. Qualitative analysis of these joins demonstrated that these junctions are indistinguishable from wild type with respect to the number of deletions. Collectively however, the SCID joins exhibited a profoundly restricted repertoire characterized by a general absence of N addition and an increase in homology mediated joining. Chapter IV provides a detailed analysis of the extent to which VDJH and κ rearrangements are found in SCID B lineage. I show that SCID mice undergo normal VDJH and kappa rearrangement at an appreciable frequency (0.1-1%), and many of these rearrangements are functional (86%, 44% respectively). About 40% of the rearrangements could have arisen by homology mediated ligation, suggesting that homology is used in the resolution of coding ends in the absence of functional SCID protein. Despite the relatively high level of Ig gene rearrangement, SCID mice do not have detectable pre-B cells. Moreover, SCID mice transgenic for either fully rearranged heavy and light chain Ig do not exhibit increased leakiness. My data provides evidence which suggests that the *scid* defect has a profound and lasting effect on B cell development beyond deficiencies in V(D)J recombination.
ACKNOWLEDGMENTS

I wish to thank my supervisor Gillian Wu for the opportunity to work in her lab. Gill's passion for science stems from an honest desire to achieve greater understanding rather than from the pursuit of personal glory - truly the most noble motive to undertake a career in science. Moreover, I am indebted to Gill for teaching me to persevere in the face of adversity.

I wish to thank my supervisory committee, Bob Phillips, Cindy Guidos and Susanna Lewis for their valuable scientific advice - this thesis is far superior for it. In particular, I am grateful to Bob and Susanna for their insightful scientific discussions and encouragement above and beyond the call of duty. Your kindness shall not be forgotten. I would also like to express my appreciation to my "deputy" thesis advisor Dr. Fred Bertrand for his assistance and kindness during the writing of this manuscript.

I would like to express my gratitude to Connaught, the National Cancer Institute of Canada and the Medical Research Council of Canada for their financial support during various stages of my training.

During the course of my studies, I have been fortunate to meet with many unique characters who have enriched my life with their friendship. I am grateful to Sue Chappel, Burshtyn, Juan Carlos Zúñiga-Plücker and Kevin Galley - their enthusiasm and love of science has truly been inspirational. I am indebted to Bobby-Joe Ray, Devila Stoddart, Alison Conner, Tania Benatar, "Doug" Michie and Heather Fleming for infusing laughter and mischief into my scholarly existence.

Finally, I wish to thank my family and my long time friend, Angela Gossmann, for their unwavering support, for never allowing me to doubt myself, and more importantly, for reminding to me laugh at the absurdities of life. I could not have done this without you.
"UT RESURGAM"

-Pennycook family motto

For Mutti, Pup and my beloved sisters "Bear", "Lupé", and "Swee".
-This thesis is a testimony to your unwavering faith in me.

I also wish to dedicate this thesis to the memory of Sandra Bridle.
- I have not forgotten your dream nor your legacy.
LIST OF ABBREVIATIONS

Ab: antibody
Ag: antigen
A-MuLV: Abelson-murine leukemia virus
bp: base pair
BM: bone marrow

C: Cytosine
Cµ, Cδ, Cγ, Cα, Cε, Cκ, Cλ: constant region domains from the isotypes µ, δ, γ, α, ε, κ, and λ respectively
CD: cluster of differentiation
CDR: complementary determining region
CSC: cleavage synaptic complex
Cys: Cysteine

D: immunoglobulin “diversity” element
DJH: rearrangement of DH element to a JH element
Dµ: protein encoded by most DJH rearrangements in reading frame 2
DNA: deoxyribonucleic acid
DNA-PK: DNA-dependent protein kinase
DNA-PKcs: catalytic subunit of DNA-PK
DSB: double strand DNA breaks
ds: double stranded
ΔVJK: aberrant-sized kappa rearrangements

EDTA: ethylenediamine tetraacetic acid
Fc: Constant region of an immunoglobulin molecule
G: Guanine

H: pertaining to the immunoglobulin heavy chain
HSA: heat stable antigen

ICE: interleukin-1 β converting enzyme
IEF: isoelectric focusing
Ig: immunoglobulin
Igα and Igβ: proteins associated with surface Ig responsible for signaling engagement of the Ig receptor

J: immunoglobulin “joining” element
κ: immunoglobulin kappa light chain
kD: kilodalton
Ku70: 70 kD DNA-binding subunit of the DNA-PK
Ku80: 80 kD DNA-binding subunit of the DNA-PK
L: pertaining to the immunoglobulin light chain
λ: immunoglobulin lambda light chain
λ5: protein homologue to the immunoglobulin lambda constant domain

mAb: monoclonal antibody
μMT: mouse homozygous for a disruption of the transmembrane region of the m heavy chain
mRNA: message ribonucleic acid

N: non-templated nucleotide addition at junctions of recombined genetic elements

P: palindromic nucleotide addition derived from end of one of the recombining elements
pBCR: pre B cell receptor
PCR: polymerase chain reaction

RAG: recombination activating gene
RF: reading frame
RSS: recombination signal sequence

SCID: severe combined immunodeficiency
SDS: sodium dodecyl sulfate
SSC: sodium citrate/saline solution

T: Thymidine
TCRα: T cell receptor alpha chain
TCRβ: T cell receptor beta chain
TCRδ: T cell receptor delta chain
TCRγ: T cell receptor gamma chain
TdT: terminal deoxyribonucleotidyl transferase

V: immunoglobulin "variable" element
VJ: rearrangement involving V and J elements, usually refers to the light chains of the Ag receptors
V(D)J: process of somatic recombination specific to the immunoglobulin and T cell receptor gene loci
LIST OF FIGURES

Figure 1-1: Clonal Selection........................................................................................................6
Figure 1-2: Schematic of Immunoglobulin Domain Structure.................................................8
Figure 1-3: Higher Order Structure of Immunoglobulin..........................................................10
Figure 1-4: Genetic Organization of the Mouse Immunoglobulin Loci.................................12
Figure 1-5: The Recombination Signal Sequence....................................................................20
Figure 1-6: "Normal" Recombination......................................................................................23
Figure 1-7: Nonstandard Recombination..................................................................................25
Figure 1-8: Mechanism of V(D)J Recombination.................................................................39
Figure 1-9: V(D)J Recombination Cleavage Reaction.............................................................44
Figure 1-10: B Cell Development............................................................................................51

Figure 2-1: Schematic Diagram of the DJH PCR Assay...........................................................94
Figure 2-2: Schematic Diagram of the VDJH PCR Assay......................................................95
Figure 2-3: Schematic Diagram of the κ PCR Assay...............................................................96

Figure 3-1: Southern Blot Analysis of DJH Rearrangements..................................................116
 in SCID BM
Figure 3-2: Sequence Analysis of C.B-17 DJH Structures.....................................................121
Figure 3-3: Sequence Analysis of SCID DJH Structures.......................................................124

Figure 4-1: Southern Blot Analysis of VDJH Rearrangements.............................................139
 in SCID BM
Figure 4-2: Southern Blot Analysis of VDJH Rearrangements ...........................................141
 in SCID Whole BM
Figure 4-3: Sequence Analysis of SCID VDJH Structures...................................................146
Figure 4-4: Southern Blot Analysis of SCID κ Rearrangements in BM...............................150
Figure 4-5: The Aberrant κ Rearrangements Arise From Homology-Mediated Resolution ........................................... 151

Figure 4-6: Sequence Analysis of SCID κ Rearrangements .............................................................................. 155

Figure 4-7: Cutting on the Wrong Side of the Heptamer .................................................................................. 164
LIST OF TABLES

TABLE 1-1: Summary of Mouse B Cell Development 52
TABLE 1-2: Summary of Mammalian DSB Repair Mutant Complementation Groups 77
TABLE 2-1: Immunoglobulin PCR Primers. 99
TABLE 2-2: PCR Conditions for Detection of Ig Gene Rearrangements 101
TABLE 3-1: Typical Densitometric Analysis of DJH Rearrangements 118
TABLE 3-2: Quantification of DJH Structures in BM and Spleen 119
TABLE 3-3: Comparison of DJH Structures Derived from SCID and C.B-17+/+ BM and BALB/c Fetal Liver 126
TABLE 3-4: Comparison of DJH Structures With and Without N or P Additions 127
TABLE 4-1: Typical Phosphorimager Analysis of VJDH Rearrangements Observed in C.B-17 and SCID Mice Whole and Enriched BM 142
TABLE 4-2: Relative Frequency of SCID VDJH Rearrangement as Compared to C.B-17 143
TABLE 4-3: Typical Phosphorimager Analysis of VJk Rearrangements in SCID and C.B-17 Mice 152
TABLE 4-4: Comparison of VJk Rearrangements in SCID Whole and Enriched BM to Wild-type C.B-17 153
TABLE OF CONTENTS

CHAPTER I: INTRODUCTION.............................................................................................................1

OBJECTIVE.......................................................................................................................................2

THE IMMUNE SYSTEM.....................................................................................................................4
CLONAL SELECTION THEORY..........................................................................................................5
IMMUNOGLOBULIN..........................................................................................................................7
  Ig STRUCTURE..............................................................................................................................7
    • Ig Domains.............................................................................................................................9
    • Ig Isotypes.............................................................................................................................11
      Light Chain Isotypes.............................................................................................................11
      Heavy Chain Isotypes..........................................................................................................13
      Isotype Function...................................................................................................................14
      The Variable Domain............................................................................................................15

THE VARIABLE REGION..............................................................................................................16
GENETICS OF THE VARIABLE REGION.......................................................................................16
CHROMOSOMAL ORGANIZATION OF THE MOUSE Ig LOCI............................................................17
  • Heavy Chain Locus................................................................................................................17
  • Light Chain Locus..................................................................................................................18
GENERATION OF DIVERSITY.........................................................................................................18
  • V(D)J Rearrangement.............................................................................................................18
BASIC MECHANISM OF V(D)J RECOMBINATION.........................................................................19
  • The Substrate........................................................................................................................19
    Recombination Signal Sequences........................................................................................19
    Products of Recombination.................................................................................................21
      Nucleotide Addition............................................................................................................22
  • Nonstandard Junctions...........................................................................................................24
THE ENZYMATIC MACHINERY........................................................................................................26
  • The Recombinase Activating Genes......................................................................................26
  • Terminal Deoxynucleotidyl Transferase...............................................................................28
  • DSB Repair Enzymes.............................................................................................................29
CONTROL OF V(D)J RECOMBINATION.........................................................................................30
  • Regulation of RAG Activity...................................................................................................30
  • Locus Accessibility................................................................................................................32

MECHANISM.....................................................................................................................................34
STAGE I RECOGNITION....................................................................................................................34
  • Mutation Analysis..................................................................................................................34
  • RSS Recognition...................................................................................................................36
STAGE II CLEAVAGE & RECOMBINATION INTERMEDIATES.........................................................38
  • In vitro Systems....................................................................................................................43
  • Synapsis...............................................................................................................................46
STAGE III RESOLUTION..................................................................................................................47
B CELL DEVELOPMENT & Ig REARRANGEMENT

ONTOGENY OF B CELL DEVELOPMENT

Ig REARRANGEMENT & B CELL DEVELOPMENT

• Pro-B Cells
• Pre-B Cells
• Immature B Cells
• Mature B Cells

ALLELIC EXCLUSION

• The Surrogate Light Chain

THE SCID MOUSE

DISCOVERY AND PHENOTYPE

• SCID and Lymphoid Development
• SCID Mice Are Leaky
• Conditions Affecting Lymphoid Development in SCID Mice

ANTIGEN RECEPTOR TRANSGENIC MICE

ADOPTIVE TRANSFER OF WILD TYPE BM OR THYMOCYTES

ABROGATION OF NK OR MONOCYTE ACTIVITY

SCID/Bcl-2 TRANSGENIC MICE

SUBLETHAL γ-RADIATION

MENAGE à TROIS: SCID, V(D)J RECOMBINATION & DSB REPAIR

THE SCID DEFECT AND V(D)J RECOMBINATION

SCID MICE ARE DEFECTIVE IN DSB REPAIR

• The Break Connecting DSB Repair and V(D)J Recombination

CHARACTERIZATION OF THE SCID DSB REPAIR MUTATION

• Characterization of DSB Repair Mutants
• The DNA-Dependent Protein Kinase
• Ku
• The SCID Defect is DNA-PKcs

CHARACTERIZATION OF THE SCID PROTEIN

• Role of DNA-PK in DSB Repair

THESIS SUMMARY
CHAPTER II: METHODS

ISOLATION AND PURIFICATION OF BONE MARROW ..... 88
CELLS AND DNA
MICE AND CELL LINES ............................................. 88
DETECTION OF SERUM IMMUNOGLOBULIN ............... 88
PURIFICATION OF CELLS ......................................... 89
DNA PREPARATION ................................................. 91
  • Phenol/Chloroform Extraction ................................. 91
  • Direct Cell Lysis .............................................. 92
  • Miniprep DNA Extraction ..................................... 92

PCR ASSAYS .......................................................... 93
D-Jh REARRANGEMENT ............................................. 93
V-DJh REARRANGEMENT .......................................... 98
K REARRANGEMENT .............................................. 100
PCR ASSAY CONDITIONS .......................................... 100

DETECTION OF Ig REARRANGEMENTS ..................... 103
SOUTHERN BLOT ANALYSIS ...................................... 103
PROBING ............................................................... 103
  • Random Primed Detection .................................. 104
    Radioactive Detection ........................................ 104
    Non-Radioactive Detection ................................. 105
    Radioactive Oligo Labeling ............................... 106
QUANTIFICATION ................................................... 107
STATISTICAL ANALYSIS .......................................... 107
CLONING AND SEQUENCING .................................... 109
CHAPTER III: HIGH FREQUENCY OF NORMAL DJH REARRANGEMENT IN SCID MICE...................................................................................................................111

INTRODUCTION.........................................................................................................................................................112

RESULTS.......................................................................................................................................................................115
QUANTIFICATION OF DJH STRUCTURES......................................................................................................................115
CHARACTERIZATION OF DJH STRUCTURES..................................................................................................................120
DJH REARRANGEMENTS IN C.B-17 BONE MARROW.................................................................................................120
DJH REARRANGEMENTS IN SCID BONE MARROW......................................................................................................123

DISCUSSION.................................................................................................................................................................129
FUNCTIONAL DJH JOINTS IN SCID LYMPHOID PROGENITORS...................................................................................129
FETAL-NATURE OF THE DJH JOINTS IN SCID MICE...............................................................................................131
IMPLICATIONS FOR THE LEAKY PHENOTYPE...........................................................................................................132
CHAPTER IV: NORMAL VDJH AND V\(\kappa\) REARRANGEMENTS IN NON-LEAKY SCID MICE

INTRODUCTION

RESULTS

VDJH REARRANGEMENTS IN SCID BONE MARROW
- High Frequency of Productive VDJH Rearrangement in SCID Mice
- Many SCID VDJH Joins Arise by Homology-Mediated Recombination
- Imprecise RSS Cleavage in SCID VDJH Structures
- SCID Mice Exhibit Increased Incidence of P Addition & Reduced N Insertions

SCID MICE UNDERGO "NORMAL" & ABERRANT \(\kappa\) REARRANGEMENT
- N Insertion in SCID \(\kappa\) Rearrangements
- Aberrant SCID \(\kappa\) Rearrangements Arise by Homology-Mediated Recombination

DISCUSSION

SCID MICE HAVE A HIGH FREQUENCY OF NORMAL IgH & \(\kappa\) REARRANGEMENTS

SELECTION OF SCID B220\(+\)sIgM\(-\) CELLS WITH VDJH REARRANGEMENTS

IMPRECISE RSS CLEAVAGE IN SCID MICE

V(D)J REARRANGEMENT IN CONTEXT OF THE MURINE SCID DEFECT
CHAPTER V: DISCUSSION........................................................................................................170

APPARENT DISCREPANCY WITH A-MuLV DATA.............................................................172
  •Primary Cells Lack Large Deletions in Their
    V(D)J Rearrangements................................................................................................172
  •Accounting for the Apparent Differences in Coding Joint
    Formation Frequency Between Primary and Transformed
    SCID Cells.....................................................................................................................174

RATIONALIZING REPERTOIRE RESTRICTION IN SCID MICE...............................176
DNA-PK, V(D)J RECOMBINATION AND B CELL DEVELOPMENT......................177
REMAINING ISSUES AND FUTURE DIRECTIONS.....................................................178

REFERENCES.......................................................................................................................183
CHAPTER I

INTRODUCTION
OBJECTIVE

One hallmark of B cell development is the assembly of the genetic elements encoding the variable exon of the antigen (Ag) receptor or immunoglobulin (Ig) molecule. This assembly process is termed V(D)J recombination. The goal of my thesis is to explore the assembly process of the Ig variable exon elements and its affects on B cell development. To attain this goal, I have chosen to study V(D)J recombination in a severe combined immunodeficiency (SCID) mouse that is mutant in this activity (as reviewed in (1), see below).

SCID mice have a defective double strand DNA break repair system that results in a profound lack of mature lymphoid cells (2-5). Lymphoid development is affected because the assembly of the variable region exon of Ag receptors during V(D)J recombination generates double strand DNA breaks (DSB). The failure to resolve these DNA lesions prevents further differentiation of the lymphoid cells, and typically results in cell death.

The scid mutation is not absolute as older mice occasionally become "leaky", exhibiting serum Ig and/or T cell activity (6, 7). Given this leaky phenotype, I wanted to determine the extent to which Ig gene rearrangement occurs in different loci, and if so, assess the nature of these gene rearrangements in order to understand how these junctions may have arisen in the presence of the scid mutation.

During my studies of Ig gene rearrangement in SCID mice, the scid gene product was identified as the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) (8-10). DNA-PK is a serine/threonine kinase that is composed of a catalytic domain and a DNA binding heterodimer called Ku70/80 (Ku), that is specific for double strand (ds) DNA ends. DNA-PKcs is activated only when the Ku heterodimer binds to DSBs, single stranded nicks in the DNA.
and DNA hairpins. DNA-PK is capable of phosphorylating a variety of transcription factors including hsp90 (11), replication protein A (12), and p53 (13), a molecule involved in cell cycle arrest. Consequently, the SCID mutation may have repercussions more far reaching than a defect in Ag receptor assembly.

The data presented in the subsequent chapters was generated prior to the above revelations with respect to the scid mutation. Nonetheless, the relatively high incidence of Ig gene rearrangement in SCID mice provides interesting insights into both recombination and its affects on B cell development.
THE IMMUNE SYSTEM

The role of the immune system is to protect an organism from invading pathogens. There are two branches of the vertebrate immune system that work in concert together: innate and acquired immunity. Innate immunity employs a variety of techniques to protect an organism from environmental agents e.g. physical barriers, acidic secretions, and phagocytic cells. These defense strategies tend to be non-specific in nature, and are not enhanced upon subsequent exposure to a particular agent.

Acquired immunity is characterized by its "memory", as increasingly effective defense mechanisms are elicited upon subsequent exposure to a foreign agent or antigen (Ag). The effects of the acquired immunity are primarily mediated by lymphocytes, B and T cells, and their secreted products, e.g. antibodies (Abs).

Both B and T cells arise from hematopoietic progenitor cells, but they develop in different organs. In adults, the sites of B and T cell lymphopoiesis are bone marrow (BM) and thymus respectively. These cell subsets exhibit different, yet complementary effector functions that serve to protect an organism. Each B cell possesses a receptor that can interact directly with foreign pathogens. Engagement of these receptors stimulates the B cell to secrete Abs. Abs isolated from an immunized individual can confer protection for a non-immune individual to a particular pathogen. This type of immunity is referred to as humoral. T cells are responsible for cell-mediated immunity; the transfer of these cells from an immunized individual alone will protect a naive individual from infection. T cells require Ag-derived peptides to be presented to them in the context of the major histocompatibility complex (MHC) proteins in order to be stimulated into action.
CLONAL SELECTION THEORY

Ehrlich was the first to propose that specific Abs exist prior to encountering Ag (as reviewed in (14, 15)). Jerne first rejected and then subsequently reinstated this concept as a key feature of the adaptive immune system. Eventually, Burnet and Talmage proposed that each lymphocyte cell is genetically committed to express Abs of distinct specificity for different foreign agents prior to Ag exposure. Thus, the clonal selection theory evolved to postulate that only cells which engage Ag with their specific receptor are selected to expand and differentiate, resulting in the eventual removal of the Ag. This principle of clonal selection has been shown to be true for both B and T cells, although the manner in which these two cell types interact with their Ag is fundamentally different, reflecting their different effector functions. During the maturation phase of an immune response, clones arise that have greater affinity for the original stimulus, and these cells can become memory cells better able to address future encounters with same pathogen. Figure 1-1 illustrates the clonal selection theory as it pertains to the B lineage. Clearly the inability to express an Ag receptor will have devastating effects on an organism’s ability to protect itself against pathogens.

Because B and T cells must address an almost infinite array of antigenic stimuli, the means by which they generate Ag receptors share common strategies. The focus of my thesis is to understanding how antigen receptors are generated in B cells and how their generation (or lack thereof) affects B cell development. Consequently, the thrust of my thesis will pertain to the generation of surface immunoglobulin.
Immune Challenge

Proliferation & Affinity Maturation

Differentiation

Plasma Cells

Memory Cells

Fig. 1-1: Clonal Selection. \( \downarrow \) denotes Ag and \( \uparrow \uparrow \) represents Ig specific for the Ag. Only B cells capable of interacting with an Ag can proliferate and undergo differentiation into Ab secreting plasma cells and/or memory B cells.
IMMUNOGLOBULIN

Ig STRUCTURE

The term immunoglobulin (Ig) refers to a heterogeneous population of molecules produced by B lineage cells (as reviewed in (14)). These proteins can exist in either a membrane bound or a secreted form. The term antibody (Ab) generally refers to the secreted form of Ig molecules and typically refers to a molecule of a known Ag specificity. All Ig molecules are bifunctional in nature. They possess a unique antigen recognition domain and one or more effector domains. Engagement of Ag with the Ig variable domain elicits a number of biological functions that contribute to the ultimate removal of the Ag.

All Ig molecules share a bilateral symmetrical core structure composed of two identical heavy (H) and two identical light (L) chains (H2L2) (see Figure 1-2). Each L chain is approximately 23 kDa, while a single H chain molecule is 55 -70 kDa depending on its class of Ig (see below). In most classes of Ig, a single L chain covalently associates with a H chain molecule via a disulfide bond (human IgA2 is a notable exception as the two IgL chains are covalently linked, and their association with the IgH chains occurs entirely by noncovalent forces (14)). The interaction between the H and L chains is further strengthened by noncovalent forces generated by hydrophobic residues. The two H chain molecules are homotypically linked by one to five disulfide bonds at the hinge region. All classes of membrane bound Ig exist in this basic heterodimeric form, however some classes of secreted Ig also form multimeric structures (see Isotypes). Membrane bound Ig possesses a short hydrophobic transmembrane region and a very short cytoplasmic tail. Surface Ig is always found in association with two cytoplasmic proteins, Ig-α and Ig-β. Secreted forms of Ig have a short secretory tail in lieu of the transmembrane region and cytoplasmic tail.
Figure 1-2: A. Schematic diagram of the higher order structure of secreted IgG1 illustrating the bilateral symmetry of Ig. B. Schematic diagram of the Ag binding site of an Ig molecule.
**Ig Domains**

Both H and L chain molecules are composed of discrete globular domains (Figure 1-3). Each domain consists of approximately 110 amino acids (AA). These domains are created by two anti-parallel β-pleated sheets made up of seven to nine strands that are linked by a single intrachain disulfide bond (see Figure 1-3). Further stability for the globular structure is achieved by non-covalent interactions between the hydrophobic AA of the inner faces of the two opposing β-pleated sheets (16). All Ig domains maintain a strikingly similar three dimensional conformation despite considerable differences in their primary AA sequence. This characteristic globular configuration is also found in a large variety of proteins; all molecules possessing this type of globular domain are members of the Ig gene superfamily.

Sequence analysis of Ig molecules of a particular class revealed that the most N-terminal domains of both the H and L chain molecules are highly variable in nature. These H and L chain variable regions are abbreviated as VH and VL respectively. The remaining Ig domains are relatively invariant in nature, especially among molecules of a particular isotype.

All L chains are made up of a single VL region and a single constant region (CL). H chains possess either three or four constant domains depending upon the class or isotype of H chain (see Isotypes). These constant domains are numbered sequentially from the amino terminus (CH1, CH2 etc.).

Within a single Ig molecule, the H and L chains are aligned in a parallel manner such that their V domains are juxtaposed to create an Ag binding site, and the CH1 and CL domains are linked (Figure 1-2). This alignment is a consequence of the interchain disulfide bridge, as well as noncovalent forces generated by the hydrophobic AA of the two combining Ig domains. The two H chains associate through interchain disulfide bonds in the hinge region located
Figure 1-3: Higher order structure of immunoglobulin. β strands are shown as wide ribbons. Adapted from (14) with permission from Daniel P. Stites, M.D., Basic and Clinical Immunology, 8th Edition, Appleton and Lange, 1994.
between the CH1 and CH2 domains, and this interaction is further strengthened by the noncovalent association of CH domains with their homotypic counterpart on the opposing H chain. The hinge region of all Ig molecules is rich in proline and cysteine residues which generates a loose secondary structure. The flexible nature of the hinge region permits movement between the two Ag binding site with respect to one another, facilitating their interaction with Ag.

- **Ig Isotypes**

  Both the H and L chains have alternative forms of constant regions resulting from distinct AA sequences which can be distinguished by differences in physical properties e.g. molecular weight, as well as function. These different isoforms of Ig polypeptides are referred to as classes or isotypes.

  **Light Chain**

  There are two major isotypes of light chain in most species, kappa (κ) and lambda (λ). These isotypes are encoded by distinct genetic loci on mouse chromosomes 6 and 16 for κ and λ respectively (see Figure 1-4). Both isotypes of L chain are approximately 23 kDa. A single Ig molecule possesses only one class of identical L chains. All κ chains have an identical constant region, whereas a single λ chain can express one of three possible constant regions. The subclasses of λ chain vary only slightly in their primary AA sequence, and can be distinguished serologically by Abs directed towards these differences. Despite the differences in constant regions between and among the L chain isotypes, all classes of L chain are functionally equivalent, and can pair with any of the H chain isotypes. Generally, only one L chain isotype is expressed during the lifespan of a B cell clone. Analysis of serum Ig and B cells revealed that 95% of mouse B cells use the κ class of L chain. This profound bias towards κ usage is not as pronounced in other species as κ is used only 60% human Ig molecules.
Figure 1-4: Genetic organization of the mouse Ig loci. Exons are represented by boxes. Introns and intervening sequences are denoted by lines. This figure is not drawn to scale.
**Heavy Chain Isotypes**

There are five different classes of H chain constant regions: \( \mu, \delta, \gamma, \varepsilon \) and \( \alpha \). The exons encoding these different constant regions are located on the same genetic loci (chromosome 12), downstream of the assembled V region exon (Figure 1-4). All constant regions of a given class or subclass have identical constant domains. Ig containing \( \mu, \delta, \gamma, \varepsilon \) and \( \alpha \) chains are called IgM, IgD, IgG, IgE and IgA respectively. In mouse, the IgG isotype has four alternative isoforms distinguished by minor differences in their constant regions sequences which are consequently reflected in their function.

The classes of Ig can differ in the number of constant domains: In mice IgM and IgE have four constant domains, while all IgG subclasses, IgA and IgD contain only three. Different isotypes are also distinguished by the number of disulfide bonds pairing the H chains of an Ig molecule, as well as the length of the hinge region.

While all membrane bound Ig classes exist in the prototypical \( \text{H}_2\text{L}_2 \) configuration, serum IgA and IgM exist in multimeric forms. Secreted IgA can exist in both dimer and trimer form, while serum IgM is found in either a pentameric or hexameric form. A structurally distinct J (joining) chain is associated with most polymeric forms of serum Ig through disulfide bonds (the hexameric form of IgM does not have a J chain). The dimeric form of IgA found in external secretions e.g. saliva, is associated with a second protein called secretory component.

Unlike the L chain, the constant region of an IgH chain may vary during the life of a B cell clone. Moreover, the different isotypes mediate different effector functions in response to Ag, particularly with respect to their secreted forms. With the exception of IgD, the expression of the different isotypes is mediated by a recombination event that involves the deletion of DNA encompassing the H
chain exon currently expressed, to a point just 5' of the exon encoding the H chain exon to be expressed. This process is known as class or isotype switching. Isotype switching is influenced in part by the lymphokines an activated B cell encounters during the course of an immune response.

**Isotype Function**

IgM is the first class of Ig expressed during the life of a B cell, and is consequently the first Ig secreted in response to Ag. Antibodies secreted during a primary immune response have relatively low affinity for their antigenic stimuli. However, the high valency of Ag binding sites generated by the multimeric secreted forms of IgM makes it ideal for dealing with multivalent antigens.

Antigen-complexed IgM, as well as most subclasses of IgG (mouse IgG1 excluded), are effective in the activation of the classical complement pathway. Activation of complement initiates a cascade of events which ultimately results in the lysis of infected cells.

Most subclasses of IgG possess the unique ability of crossing the placenta, thereby conferring immunity to a fetus. IgG can act to facilitate phagocytosis of microorganisms by opsonization. Cells coated with IgG, IgA or IgE target large microorganisms or infected cells for elimination through a process called antibody-dependent cell-mediated cytotoxicity (ADCC). ADCC involves the recognition of constant region of the bound Ig by Fc receptors on natural killer (NK), cytotoxic T cells and eosinophils. The engagement of these Fc receptors activates these cells, causing the release of various cytokines and granule proteins that culminate in the destruction of the cell. IgE is responsible for allergic disorders by activating basophils and mast cells to release their histamine, leukotriene and other pharmacologically active compounds.
The secreted form of IgD is very sensitive to proteolytic cleavage, and makes up only less than 0.004% of the serum Ig. Consequently, secreted IgD's role in antigen clearance appears negligible. IgD is coexpressed with IgM on mature B cells by alternative splicing of a single mRNA molecule, and its expression of IgD stops shortly after B cell activation. Mice deficient for membrane IgD expression undergo normal B cell development, and are fully capable of responding to Ag (17, 18). However, affinity maturation in these animals is delayed compared to normal mice, implicating IgD in a role for augmenting affinity maturation during the early phases of an Ag response.

The Variable Domain

The variable domains determine the antigen specificity of an Ig molecule. The first evidence demonstrating that the variable domain is responsible for Ag specificity came from experiments using proteolytic enzymes that cleave the Ig molecule at the hinge region. Only fragments bearing the amino terminal V domains of the Ig molecule are capable of binding Ag.

Sequence analysis revealed three regions of hypervariability of nine to twelve AA in length, interspersed amongst relatively conserved regions called framework regions (FR) (see Figure 1-2). Sequence analysis, and later X-ray crystallography confirmed that the characteristic Ig folds or β-sheets are mediated predominantly by the more conserved FR regions (Figure 1-3). Moreover, the hypervariable regions of the V domains form the loops attaching the β-strands of the Ig motif and project outwards. The juxtaposition of VL and VH domains in a complete Ig molecule creates a three dimensional antigen binding pocket. Because the antigen binding site is believed to be complementary to the three-dimensional surface of a bound antigen, the hypervariable regions are referred to as complementarity determining regions (CDRs). The CDRs and FR regions are numbered sequentially from the N-
terminus (Figure 1-2).

THE VARIABLE REGION

GENETICS OF THE VARIABLE REGION

The exon encoding the variable domain is assembled from germline genetic segments located on the same locus by a process known as V(D)J recombination. There are three classes of genetic elements, variable (V), diversity (D) and joining (J). Variable region exons of IgL, TCR α and γ are encoded by V and J elements only, whereas IgH, TCR β and δ variable exons also include D elements. Rearrangement generally occurs between elements located on the same chromosome. Elements belonging to different Ig (and TCR) loci are distinguished from one another by a chain designation e.g. VH, Vk etc.

Each V element encodes a short leader peptide required for transport through the endoplasmic reticulum. A short intron separates the leader peptide from the remainder of the V segment which encodes the majority of the variable exon. Light chain V segments contribute approximately 95 AA to the exon, while heavy chain V segments donate about 101 AA. V elements encode FR1-3 and CDRs 1 and 2 entirely, and partially contribute to CDR3. Promoter elements required for transcription of the assembled variable exon are located upstream of each V segment.

D elements are only found in the heavy chains of antigen receptors (i.e. IgH, TCR β and δ) of most vertebrate species. D elements have no fixed reading frame, unlike the V and J elements, and they can contribute between two to nine AA to the CDR3 of a VH domain.

J segments contribute 12-17 AA of a variable domain, which includes the remainder of CDR3 and all of FR4. An mRNA donor splice site is located
immediately downstream of each J segment, which permits splicing of an assembled V exon to a constant domain exon.

CHROMOSOMAL ORGANIZATION OF THE MOUSE Ig LOCUS

Mice possess three unlinked Ig loci which encode the Ig H chain, and the κ and λ light chains. The IgH loci resides on chromosome 12, whereas the light chain loci are on chromosomes 6 and 16 for κ and λ respectively (as reviewed in (19)).

• Heavy Chain Locus

The IgH locus exists in an extended configuration whereby multiple members of the same class of genetic element are clustered (see Figure 1-4). Members of the three variable gene element classes are located upstream of the exons encoding the various isotypes of the IgH chain (reviewed in (19)).

The mouse IgH locus is estimated to have approximately 120 different VH elements (as reviewed in (20)). These elements are subdivided into families based on sequence identity. All members of a single VH gene family share greater than 80% nucleotide identity.

There are 15 known DH elements that can be categorized into four families; two of which contain more than one member (21-26). DH elements are flanked on either side with signal sequences that target recombination, and can be thus assembled by inversion as well as deletion (see Basic Recombination). Moreover, since DH elements lack a fixed RF, rearrangement events occurring in any frame can theoretically contribute to productive IgH joins.

There are only four JH segments (19). The four JH elements are arranged in the same transcription orientation as the CH exons.
Light Chain Loci

The κ locus is also organized in an extended configuration. There are an estimated 140 Vk elements that can be grouped into eight gene families (27). The five Jk elements are clustered together 2.4 kb upstream of the sole Ck exon. Jk3 is a pseudogene because it contains a mutations in both its recombination signal sequence, as well as its RNA splice site. An estimated 50% of Vk elements are in an inverted transcriptional orientation relative to the Jk and Ck.

The λ locus is considerably less complex than either the IgH or κ loci with respect the number of different V elements. There are three Vλ, four Jλ elements, and four separate Cλ exons. This locus is organized in a clustered Vλ/Jλ-Cλ fashion (see Figure 1-4), by which rearrangement of elements are confined mostly between elements upstream of a Cλ exon. Rearrangements involving Jλ4/ Cλ4 are not found in Ig molecules because of mutations which prevent both rearrangement and splicing (19). All λ gene segments are organized in the same transcriptional orientation.

GENERATION OF DIVERSITY

V(D)J Rearrangement

The vast array of antigenic specificities of B cells is generated by a number of mechanisms. These mechanisms include germline diversity, combinatorial diversity, and junctional diversity (as reviewed in (19)).

Germline diversity refers to the large numbers of distinct genetic elements of a segment class available for recombination. For example, the IgH locus has an estimated 200 VH elements, 15 different DH elements and four JH elements.

Combinatorial diversity is generated by a process referred to as V(D)J recombination. Any element of one genetic element type can recombine to any member of its partner class. i.e. any of the estimated 200 VH elements can
rearrange to any of the 15 different DH elements, which in turn can recombine to any of the available four JH elements. This assembly of genetic elements is typically confined to developing lymphocytes (as reviewed in (28)). Secondly, combinatorial diversity is further enhanced by the combining of any resulting IgH chain with nearly any light chain.

Junctional diversity is a consequence of the imprecise nature of the V(D)J recombination reaction. Nucleotides are generally deleted from and/or added to the terminal ends of the recombining elements. This variability at the junctions of recombining elements underlies the extreme heterogeneity characteristic of CDR3. However, this diversity occurs at a cost of generating many out of frame rearrangements or joins bearing premature stop codons because the number of bases added and deleted appears to occur randomly.

**BASIC MECHANISM OF V(D)J RECOMBINATION**

*The Substrate*

**Recombination Signal Sequences**

The first insight into the mechanism underlying V region assembly came from the observation that conserved sequences flank either one or both sides of the germline sequences of the elements to be recombined in B cells. Similar signal sequences were later found adjacent to the T cell receptor (TCR) variable gene elements (28), and are referred to as recombination signal sequences (RSSs).

The RSSs are tripartite in nature, consisting of a conserved heptamer and nonamer separated by a relatively nonconserved spacer sequence of either 12 or 23 bp (see Figure 1-5A) (as reviewed in (28)). Recombination occurs only between elements with RSSs of differing spacer length. This requirement is referred to as the 12/23 rule. Not surprisingly, the organization of the Ig and TCR loci have evolved in response to the 12/23 rule, i.e. all gene segments of the
Figure 1-5: The Recombination Signal Sequence (RSS). A. Consensus RSS indicating the degree to which particular residues are conserved (29). B. Schematic diagram depicting the configuration of RSSs in the various Ig loci. Closed triangle depict 23-bp RSSs, and open triangles represent 12-bp RSSs. This diagram is not drawn to scale.
same class (V, D or J) have the same configuration of RSSs, while their joining partners are flanked with the complementary RSS (see Figure 1-5B). Thus, the 12/23 rule acts to increase the likelihood that a functional V domain will be produced.

Studies using artificial recombination substrates have revealed that a pair of complementary RSSs are both necessary and sufficient to direct recombination in cells active for V(D)J rearrangement (30-32). Replacement of coding segments by random DNA sequences still allows recombination to occur at the signal border. The optimal or consensus RSS is shown in Figure 1-5. The degree to which a residue is conserved reflects its importance in directing recombination (30-34).

The use of artificial substrates has greatly expanded our understanding of the recombination process without the complication of antigen selection, and is discussed in further detail below.

The Products of Recombination

Recombination occurs between the border of the RSS heptamer and coding element. Generally V(D)J rearrangement results in the formation of two novel DNA structures, the fusion of two coding elements, termed coding joint, and heptamer to heptamer fusion of the two RSSs to form a signal junction (see Figure 1-6). The orientation of the two RSS with respect to one another determines whether both junctions will be retained on a chromosome. As shown in Figure 1-6, RSSs arranged in the same direction will result in an inversion event whereby both the signal and coding junctions are maintained. Deletion occurs when RSS are organized in an opposite orientation. The most common configuration of RSSs in the Ig and TCR loci is the deletional orientation such that the coding junction is retained on the chromosome. However, the κ locus is estimated to have half its V element RSSs inverted with respect to the Jκ (35).
Signal junctions isolated from endogenous loci, as well as recombination substrates are characterized by the precise ligation of the RSSs at their heptameric borders (36). The loss of nucleotides from these RSSs is exceptionally rare in wild type cells, although N addition has been observed occasionally (36, 37). In contrast, coding junctions are characteristically imprecise, exhibiting a small number of nucleotide deletions and insertions. This imprecision greatly contributes to the diversity of the immune repertoire. However, this heterogeneity is achieved at the expense of generating numerous nonproductive rearrangements. The differential resolution of V(D)J recombination products is further accentuated in SCID mice whose ability to resolve coding joints is drastically compromised, while the formation of signal junctions is relatively unaffected (see SCID) (38, 39).

**Nucleotide Addition**

There are two types of nucleotide addition, non-templated (N) and templated addition. N additions are random in nature, but exhibit a preference for G/C addition (see Figure 1-8) (28). N addition is mediated by terminal deoxynucleotidyl transferase (TdT), and is discussed in more detail below.

Templated additions are only observed adjacent to coding elements that escaped nucleotide deletion (40, 41). The templated additions are complementary in reverse order to the bordering coding element residues, forming a palindrome to the bordering nucleotide sequence. As a consequence, template additions are referred to as P insertions. The incidence of these additions led to the suggestion that at least some coding elements passed through a hairpin intermediate prior to joining (see MECHANISM) (40).

Rearrangements in wild type cells generally exhibit at most four P nucleotides (42, 43), however much longer P additions (12 nucleotides) have been observed in a proportion of the rare SCID coding junctions (44-47).
A. Deletional Recombination

Figure 1-6: "Normal" Recombination. A. Deletional Recombination. B. Inversional Recombination. Symbols are as described in Figure 1-5. Processing of the coding ends is depicted by $\Xi$. 
Nonstandard Junctions

Analysis of products derived from endogenous κ loci and recombination substrates demonstrated that the V(D)J recombination process could be promiscuous in that nonstandard junctions were isolated (48-50). Nonstandard junctions include hybrid joint and open and shut junctions. It is interesting to note that this differential handling of the coding and signal ends is also observed in nonstandard recombination products, implying that the same resolution processes are in place for these aberrations in the recombination process.

Hybrid junctions involve the joining of a coding element to the RSS of its partner recombining element (Figure 1-7). Hybrid joint formation occurs often in extrachromosomal substrates, accounting for up to 30% of the total recombinants (48). These joins have also been observed in endogenous rearrangements (51, 52).

Open and shut junctions involve a cut and subsequent religation of a RSS with the coding element it originally flanked (48). A high frequency of open and shut junctions occurs at the 5' RSS of TCR DJδ rearrangements of wild type mice (53). In recombination substrates containing a pair of complementary RSSs, open and shut junctions are frequently found at only one RSS (50). Moreover, the frequency of open and shut junctions in substrates containing a single RSS is comparable to that observed for plasmids containing two complementary RSSs (50). This observation suggests that RSS recognition, cleavage and subsequent resolution of the DNA breaks can occur at a single RSS independent of synapsis. Isolation of RSS ends from substrates containing only one RSS provides support for this hypothesis (54).
Figure 1-7: Nonstandard Recombination Products. A. Hybrid Joint. B. Open & Shut Junctions. The open & shut joins are depicted as occurring at both RSSs, however open & shut joins often occur at only one RSS (see text for details). Symbols are as described in Figure 1-5.
THE ENZYMATIC MACHINERY

The process of V(D)J recombination is complex and its activity is normally confined to progenitor lymphoid cells in vivo. Despite this lineage restriction of activity, only a subset of the factors governing V(D)J recombination are specific to lymphoid cells. The other components exhibit ubiquitous expression and appear to be involved in general DNA repair pathways. Interestingly, some reaction-specific factors are not essential for recombination.

- The Recombinase Activating Genes

The recombination activating genes (RAG-1 and RAG-2), were first identified by their capacity to support recombination of exogenous recombination substrates introduced into nonlymphoid cell lines (55-57). The recombined products generated by transfecting the RAG genes into fibroblasts are comparable to those formed in lymphoid cells with one notable exception - the coding junctions lack N insertions. This lack of N addition in these transfection studies reflects the lack of TdT expression in fibroblasts (see Terminal Deoxynucleotidyl Transferase below). Importantly, both RAG genes must be expressed in order to support V(D)J recombination (57-59). Indeed, only cells actively undergoing V(D)J recombination coexpress RAG-1 and RAG-2 (57) (see B CELL DEVELOPMENT & Ig REARRANGEMENT). Moreover, mice deficient for either RAG-1 or RAG-2 are devoid of V(D)J recombination activity, and are consequently immunodeficient (58, 59). Unlike SCID mice, RAG-deficient lymphoid cells do not even initiate V(D)J recombination (60).

Collectively, these data suggested that the RAG proteins either specifically activate the recombination machinery or that they directly participate in the recombination reaction.

The genes encoding the RAG proteins are closely linked, separated by a mere eight kb in the mouse genome (57). In the mouse, both RAG-1 and RAG-2
proteins are encoded by a single exon which includes the 3' untranslated region. In the mouse RAG-1 and RAG-2 proteins are 1040 and 527 AA in length respectively (56, 57). There is no homology between the two RAG proteins (57). However, RAG-1 and RAG-2 proteins are each highly conserved throughout vertebrate evolution (61).

Neither RAG protein bears any significant homology to any know protein (56, 57). However, three sequence motifs were identified in RAG-1 the provided possible insights into its function. These motifs included a nuclear localization signal, a zinc-finger-like Cys-His motif, and a region within the carboxy the carboxy terminal half of RAG-1 protein that shares limited sequence similarity with the yeast HPR1 gene product (62). The last motif created some excitement because the region of resemblance to the HPR1 gene in turn shares some sequence identity with *Saccharomyces cerevisiae* DNA topoisomerase I (63). However, mutations in RAG-1's putative topoisomerase active-site tyrosine do not affect either the frequency or the quality of recombination suggesting the putative topoisomerase motif lacks functional significance in V(D)J recombination (64). The putative RAG-1 nuclear localization signal and zinc finger motif in the N-terminal domain are also not essential for activity, as accurate V(D)J recombination activity is still observed upon deletion of residues 15-383 (65-67). However deletion of residues 994 to 1040 within the C-terminal domain abolish V(D)J recombination activity (65). These analyses suggest that the critical regions for V(D)J recombination lie within the C terminal domain of RAG-1. This recombinationally active truncated RAG-1 protein containing residues 344-1008 is referred to as core RAG-1 protein (65, 66).

Unlike RAG-1, the N-terminal domain of RAG-2 is critical for V(D)J activity. Removal of up to 25% of the C-terminal domain in RAG-2 still permits V(D)J recombination (65, 68, 69). This dispensible C-terminal region includes a highly
conserved acidic region which is a common motif in transcription factors (57). Thus RAG-2's role in V(D)J recombination is not likely to be as transcription factor. Surprisingly, mutations in the less conserved N-terminal regions have a more profound affect on V(D)J recombination activity (68, 69). RAG-2 AA residues 1-381 are the minimal requirement for recombination; this protein is subsequently referred to as core RAG-2 (65, 68, 69).

Co-expression of the core RAG proteins is sufficient to support V(D)J recombination in nonlymphoid cells (67). In fact, there appears to be a slight increase (two to three fold) in recombination activity over full length RAG proteins (67). These core proteins are much more soluble than the native RAG proteins.

**Terminal Deoxynucleotidyl Transferase**

Terminal deoxynucleotidyl transferase (TdT) is a lymphoid specific enzyme responsible for generating the majority of N additions found in recombination products (70-72). Indeed, coexpression of TdT with the RAG proteins in nonlymphoid cell lines restores the incidence of N addition found in recombination products (72). In general, the frequency of N addition correlates well with the level of TdT expression (73, 74). As discussed earlier, N additions are found predominantly in coding junctions, however, they have been observed in signal junctions on occasion (37). The incidence of signal junction N addition is associated with high expression of TdT protein (37).

TdT activity is not essential for V(D)J recombination. First, TdT expression is developmentally regulated. During B cell development, TdT expression is terminated during the stage at which most light chain rearrangements occur (see Table 2-1 (75-79)). Moreover, TdT is not expressed during fetal development (75), a phenomenon observed in a most vertebrate species (as reviewed in (28)).
Secondly, TdT deficient mice generate B cells virtually devoid of non-templated nucleotide addition at their coding junctions (70, 71).

The purpose underlying TdT's dichotomous expression during ontogeny is the subject of much debate. The ensuing restricted fetal repertoire has been proposed to be necessary for conferring special properties to the immune system such as the establishment of polyspecific Ag receptors to enhance protection of the neonate ((28) and references therein). Surprisingly, both TdT-deficient mice and mice expressing TdT transgenes during fetal development are capable of efficient immune responses (80, 81), thus casting doubt of the necessity of N addition for effective B and T cell responses.

Nonetheless, the primary function of TdT appears to be the diversification of the immune repertoire by increasing the length of the CDR3. In the absence of TdT expression, junctions routinely occur at regions of homology resulting in a severe restriction in the repertoire (26, 71, 80, 82-85). The presence of terminal homologies in coding elements does not enhance the efficiency of resolution in recombination substrates, but rather only affects the quality of the resolved junction (85-87). Interestingly, coding junctions without N addition that arise in the presence of TdT exhibit a paucity of homology-directed recombination (70, 84, 86, 87). This observation suggests that the mere presence of TdT interferes with the mechanism(s) underlying homology-mediated coding end resolution.

**DSB Repair Enzymes**

The SCID mouse provided the first link between V(D)J recombination and DNA repair mechanisms. The profound deficiency in lymphoid development in SCID mice is a consequence of an inability to resolve coding ends (2, 38, 88). This discovery sparked an intensive effort to identify other mutants in DSB repair and V(D)J recombination. These findings will be discussed in detail below (see SCID).
CONTROL OF V(D)J RECOMBINATION

The process of V(D)J recombination is mediated by an elaborate series of mechanisms that intersect with those governing development of early B and T lymphoid cells (see B CELL DEVELOPMENT AND Ig REARRANGEMENT). Moreover, V(D)J recombination requires introduction of DSBs. Failure to resolve such DSBs can result in genomic instability, while misappropriate joining can contribute to the pathogenesis underlying lymphoid malignancies (89). Consequently, the mechanisms governing V(D)J recombination must encompass devices that monitor accessibility of different genetic loci, as well as safeguards to maintain the genomic integrity and minimize oncogenic translocations.

Regulation of RAG Activity

The lineage and developmental stage specificity of V(D)J recombination is controlled in part by the co-expression of the RAG genes (56, 57). Indeed, failure to express either RAG protein is equated with an inability to initiate V(D)J recombination and arrest in lymphoid development (58-60) (see MECHANISM). Moreover, studies of cell lines transfected with heat-shock inducible RAG genes reveal that V(D)J recombination is only observed under conditions promoting RAG gene expression (90, 91).

RAG gene expression fluctuates during B cell development (see B CELL DEVELOPMENT & Ig REARRANGEMENT) (92). High levels of RAG transcription are observed in pro-B/pre-BI and small pre-BII cells (77, 92), where IgH and IgL rearrangement respectively is occurring (77, 93, 94). However, RAG transcription is not observed in large cycling pre-BII cells which express the pre-B cell receptor (pBCR) and give rise to small pre-BII cells (92). Thus, the two peaks of RAG transcription are separated by a proliferative expansion phase. This downregulation of RAG transcription is not merely a function of cell cycle
status, as pro-B/pre-BI cells enriched for cells in S/G2 and M exhibit high levels of RAG transcription. RAG gene transcription is again terminated upon differentiation into sIgM+ immature resting B cells (77, 92).

During B cell differentiation, the minimal three V(D)J rearrangements required for Ig expression occur over several cell divisions (95). The presence of unresolved DSBs during S phase or mitosis (M phase) is deleterious to proliferating cells. Thus, the requisite introduction of DSBs and the involvement of a DSB repair pathway suggested that the process of V(D)J recombination may be restricted to a particular cell cycle phase (96). One manner in which this may be achieved is through post-translational modifications of either of the RAG proteins.

Cell cycle analysis of RAG protein levels in pro-B/pre-BI cells and thymocytes revealed that RAG-2 protein accumulates in cells residing in the G0/G1 phase of cell cycle, but is virtually undetectable during S and G2/M phase (92, 97). RAG-1 protein levels however, fluctuate only moderately over the cell cycle (98). Interestingly, the half life of RAG-1 in RAG-2 deficient mice is increased over ten-fold, suggesting RAG-1 association with RAG-2 targets it for degradation, perhaps during the transition from G1 to S phase (99). Furthermore, DSBs at the RSS are almost exclusively found in G0/G1 cells (60), a finding consistent with V(D)J recombination activity being restricted to this phase of cell cycle.

The fluctuation in RAG-2 protein levels during cell cycle is not reflected at the transcriptional level in pro-B/pre-BI cells indicating that posttranscriptional modifications are likely responsible for its periodicity (92, 97, 100). RAG-2 protein stability is governed by the two regions lying within 89 AA of the carboxy-terminal (100). These motifs include the cyclin dependent kinase (CDK) phosphorylation site at Thr490Pro491 and a cationic stretch between residues
499-508. (It is noteworthy that this region is not essential for V(D)J recombination activity (65, 68, 69)). However, mutation in either of these motifs prevents RAG-2 degradation at the G1/S boundary in cell cycle, resulting in an accumulation of RAG-2 protein throughout cell cycle in transfected fibroblasts (100).

Furthermore, restriction of recombination intermediates to G0/G1 is abolished in RAG-2 deficient mice transgenic for RAG-2 protein with an altered CDK phosphorylation site, thus indicating that these motifs couple V(D)J recombination to cell cycle.

- **Locus Accessibility**

The expression of the RAG proteins alone is not sufficient to permit rearrangement of the endogenous Ig or TCR loci of non-lymphoid cells (57). TCR genes are generally not rearranged in B cells, nor are Ig gene rearranged in T cells. Moreover, recombination is generally temporally and developmentally regulated within a lineage. IgH rearrangement in mice typically precedes light chain rearrangement, and D to JH joining invariably occurs prior to VH to DJH joining. (V to DH joining is rarely found without prior DJH joining). However, DJβ rearrangements are occasionally observed in B cells and DJH joints are observed in T cells (21, 101, 102). The significance of these lineage inappropriate rearrangements is not known.

This temporal regulation of V(D)J recombination is believed to be achieved by locus accessibility (103). Accessibility of a locus is measured by three different criteria: germline transcription, DNase I sensitivity and methylation.

Numerous studies have demonstrated transcription of Ig and TCR loci to be rearranged prior to, or concurrent with recombination of the locus (as reviewed in (28). The transcripts detected include the unrearranged J-Cμ regions, as well as transcription of the separate variable elements to be recombined. Indeed, such germline IgH transcripts are readily detectable in RAG deficient mice (59). There
is a high correlation between germline transcription and recombination of a particular locus in numerous inducible systems (90, 104-106). Moreover, germline transcription of loci not rearranged at a particular stage of development is generally absent. For example germline κ transcripts are usually not observed in pro-B cells (105) (see Ig Rearrangement and B Cell Development).

The link between transcription and recombination is not absolute. There have been numerous reports of rearrangement in the absence of detectable transcription (as reviewed in (28)). Moreover, exogenous recombination substrates whose level of transcription can be varied over four orders of magnitude have unaltered recombination frequencies (107). These data suggest that transcription likely represents only one of several ways to activate the endogenous substrate, or is merely an indirect means of determining activation of locus accessibility (28).

Nuclease sensitivity is another means of measuring the chromatin status of a particular locus. Ig loci enhancers and various segments of accessible loci are hypersensitive to DNase digestion (108, 109). This hypersensitivity is found in recombination substrates. Moreover, a substrate containing germline TCR V-,D- and J-β segments that undergoes DJβ rearrangement when introduced into a pre-B cell line, is DNase hypersensitive at the Dβ and Jβ elements (110, 111).

Hypomethylation correlates well with transcriptionally active loci, denoting an open chromatin configuration (112, 113). Hypomethylation of transgenic substrate V gene segments is associated with efficient rearrangement (114-116). The level of methylation in plasmid recombination substrates also correlates with efficient rearrangement (107).

However, the hypomethylation of substrates alone is not sufficient to guarantee recombination. The endogenous JH locus of B cells with a neomycin resistance gene inserted upstream of the Eμ element is hypomethylated, but
remains in germline configuration (117). Moreover, plasmid recombination substrates unable to replicate in eukaryotic cells can recombine whether they are methylated or not. Replicated plasmids did not undergo recombination suggesting that methylation itself does not prevent access of the V(D)J recombinase, but likely primes the chromatin to convert into an inaccessible state after replication (107).

All told, the mechanisms governing accessibility, and indeed the understanding of the chromatin architecture remains to be elucidated. It is likely that the accessible recombination state can be induced by different means at different stages of B cell development. It remains to be determined whether and how such different mechanisms of accessibility affect the outcome of recombination.

MECHANISM

The process of V(D)J recombination can be separated into three basic phases: RSS recognition, DNA cleavage and joining of the broken ends. Initially, understanding of the mechanisms governing recombination was limited to analysis of the effects of RSS mutations on recombination, as well as analysis of the recombination end products. Our understanding of these three stages has increased substantially in the past few years.

STAGE I RECOGNITION

• Mutation Analysis

Mutational analysis of RSSs in exogenous recombination substrates provided a means of assessing which residues are responsible for targeting rearrangement without the bias incurred by cellular selection. Both the
conserved heptamer and nonamer motifs play critical roles in targeting the recombinase. RSS substrates that contain a heptamerless RSS do not undergo complete recombination at a measurable frequency (32). Absence of a nonamer in one of the RSS pairs also greatly diminishes the frequency of recombination to about 1% of consensus RSSs (32, 33). The ability of these nonamerless substrates to sustain a small degree of recombination is thought to reflect the promiscuous nature of recombination that permits nontraditional recombination events such as V gene replacement (29, 61). Many VH genes (87%) contain an embedded heptamer near their 3' end of the coding sequence (118). Thus complete IgH genes bearing an embedded heptamer can undergo a subsequent rearrangement event with an upstream VH element containing an intact RSS (119, 120). It is believed that this process functions to replace potentially harmful rearrangements, although direct evidence for its role is lacking.

Mutations in either of the first three residues of the heptamer renders these recombination substrates inactive (30, 32, 33). However, alterations in the nonamer-proximal residues have a much less dramatic effect, resulting in recombination frequencies 70-90% of consensus RSSs (32). Interestingly, the quality of the resulting signal junctions is indistinguishable from wild type RSSs, i.e. the wildtype and mutant RSSs are joined precisely in a head to head fashion (33). This finding suggests that while these nonamer proximal mutations reduce the joining efficiency of the RSSs, they do not influence the site specificity of the reaction (33).

Mutations at position six or seven in the nonamer's AT rich tract result in a significant reduction in recombination in transient assays (32). Chromosomally integrated recombination substrates are more sensitive to alterations at position five (33). The discrepancy between the exogenous and integrated substrates may reflect slight differences in tertiary structure resulting from differential
chromatin association between episomal and integrated substrates (33). Double mutations within the poly-A tract have a more drastic effect on recombination frequency, resulting in recombination frequencies comparable to nonamerless RSSs. Furthermore, alterations of residues that normally flank the nonamer's AT core to A reduces recombination to less than two percent of consensus RSSs (33). All together, these results suggest that efficient recombination requires at least three consecutive A residues, and no more than five within the nonamer in order to efficiently bind and accurately assess the distance between the nonamer and heptamer (33).

The spacer sequence is poorly conserved, however its length of either 12 or 23 bp is strictly maintained (30, 32). Alterations of RSS spacer length by more than one bp reduces recombination frequencies to the level observed in heptamerless RSS substrates (32, 33). Replacement of the entire spacer sequence with completely random nucleotides did not result in a significant alteration in recombination frequencies (30, 32, 33), consequently, the spacer sequence was believed to be relatively inert in influencing the recombination process. However, these studies involved comparisons of separate substrate transfections, and lacked internal controls for the reaction. Moreover, alignment analysis of over 300 RSSs, revealed sites of conservation within the spacer sequence, particularly position five of the spacer (34, 121, 122). Using competitive recombination substrates to assess the influence these conserved residues, it was found that RSSs lacking the conserved residue were used only 85% of the wild type counterpart (34), indicating that the spacer does indeed influence the frequency of recombination.

- **RSS Recognition**

Headway into understanding how recombinase recognizes RSSs is finally being made. The inability of RAG-deficient mice to initiate RSS cleavage,
suggested that the RAG proteins functioned in the initial RSS recognition and cleavage stage of recombination (60). However, it was also possible that the RAG proteins functioned to activate the enzymatic components necessary for recombination. The first indication that RAG-1 protein might directly be involved in RSS recognition come from the identification of RAG-1 mutants that are hypersensitive to alterations in the coding sequences directly adjacent to the RSS (123, 124). The establishment of an in vivo RSS specific cleavage assay where only the core RAG proteins are required for RSS cleavage provided further supporting evidence that the RAG proteins are directly involved in RSS recognition and cleavage (125) (see stage II).

The role of the different RSS motifs in directing recombination is finally being elucidated. Several findings suggest that the RAG proteins initially bind the nonamer, while the heptamer is key in directing accurate cleavage and coding end hairpin formation. (1) Substrates bearing only a nonamer motif can successfully compete with a consensus RSS for RAG protein recognition in the in vitro cleavage system (126). (2) Heptamerless substrates were subjected to inaccurate nicking centering approximately in the region where the heptamer generally lies, however no hairpin intermediates were isolated (126, 127). (3) Cleavage of oligo substrates containing only a heptamer motif alone showed a five to ten-fold reduction in the frequency of nicking and hairpin formation (see stage II, Figure 1-9)(126, 127). (4) Mutations in the first three residues of the heptamer result in reduced, as well as inaccurate nicking comparable to substrates lacking the entire heptamer. Moreover, coding end hairpin formation is not observed even with pre-nicked substrates, suggesting that these residues are critical for complete cleavage.

Using plasmon resonance and an in vivo one-hybrid DNA binding assay, RAG-1 alone could mediate RSS recognition (128, 129). However, RAG-2 alone is
unable to associate with the RSS. The nonamer appears to be the key element in the initial RSS recognition by RAG-1, however RAG-1 also possesses a low affinity for the heptamer alone. RAG-2 can only associate with the RSS in the presence of RAG-1, suggesting that the nonamer/RAG-1 association recruits RAG-2. Moreover, RAG-1 association with the heptamer is enhanced upon RAG-2 recruitment. The anchoring of RAG-1 to the nonamer and subsequent recruitment of RAG-2 leads to cleavage at the heptamer/coding element border. The exact stochiometry of this reaction, particularly with respect to the 12/23 rule has yet to be determined.

STAGE II CLEAVAGE AND RECOMBINATION INTERMEDIATES

For many years, the enzymatic processes underlying V(D)J recombination could only be inferred from the analysis of its end products made in cells. The precision with which the signal ends are joined suggested that the cleavage occurred between the 5' end of the heptamer and the adjacent coding segment. Moreover, unlike N addition, P insertions are only found at coding junctions suggesting that coding ends may be processed differently than signal ends.

Isolation of DSB at the signal border of the TCR δ locus provided the first physical evidence of possible recombination intermediates (130) (see Figure 1-8). These ends were first detected by restriction enzyme digestion analysis of thymocyte DNA. The sensitivity of signal ends to exonucleases provided direct evidence that they exist as DSBs. Similar signal end intermediates derived from the Ig heavy and κ chain loci and TCR β loci were subsequently found, indicating that these structures are universal to rearranging loci, and not unique to the TCRδ locus (60, 131).

Fine-structural analysis into the nature of the signal ends was made possible by linker mediated PCR (LM-PCR). Briefly, LM-PCR involves the ligation of a
Figure 1-8: Mechanism of V(D)J Recombination. P insertions are boxed, N additions are in lower case.
linker which is blunt at one end, and staggered at the other such that only blunt ended DSBs are ligated. This technique requires that a 5' phosphate be present on the DSB in order for ligation to occur. Ig or TCR specific DSBs are detected by PCR amplification using primers specific for the linker and the locus in question. Using this technique, V(D)J recombination mediated DSBs were determined to be blunt, and 5' phosphorylated terminating precisely at the 5' heptamer border (60, 132). Neither treatment of the DNA with agents that convert single strand extensions into blunt ends nor the use of modified linkers with short 3' or 5' extensions significantly enhance the level of signal end detection, confirming that most signal end intermediates are predominantly blunt in nature.

Interestingly, these blunt RSS ends accumulate preferentially during G0/G1 phase of cell cycle (60), suggesting that initiation of V(D)J recombination occurs during this stage of cell cycle. However, accumulation of V(D)J recombination RSS DSBs is dependent not only upon their generation, but also their resolution. Thus, this finding is insufficient to definitively establish that V(D)J recombination in initiated in G0/G1. Nonetheless, two other findings indicate V(D)J recombination is initiated at this stage of development. First, resting cells exhibit a slight elevation in steady state transcription of the RAG genes (92). Secondly, RAG-2 protein is only observed in transformed lymphoid cell lines, thymocytes and pro-B/pre-BII cells in G0/G1 (92, 97, 98). Moreover, in a study using a temperature sensitive A-MuLV line, signal end intermediates were only observed in conditions that inhibited cell division, but promoted RAG activity (131). Moreover, resolution of these RSS ends was coincident with restoration of the cells conditions promoting proliferation, thus indicating that efficient signal joint formation may require cell cycle progression.
Coding end intermediates were first detected in thymocytes of SCID mice, but have not been observed in lymphocytes of normal mice (88, 130). As discussed earlier and in depth below (see SCID), SCID mice have a defective DSB repair pathway, which interferes with efficient coding joint formation resulting in a profound deficiency of mature B and T cells (see SCID). Thus, the isolation of unresolved coding end intermediates provided a major breakthrough in our understanding of the molecular process of V(D)J recombination.

Unlike signal end DSBs, the SCID coding end intermediates were resistant to exonuclease digestion suggesting that these ends existed in a modified state. The nature of this modified state was established using two-dimensional (2-D) gel electrophoresis (88). In this instance, the first dimension is non-denaturing; the second denaturing. DNA separation is on the basis of size in both directions, but if DNA is in a "hairpin" configuration, it will run at twice its size under denaturing conditions, migrating off the diagonal. Coding end intermediates, unlike the signal DSBs, migrate off the diagonal when subjected to 2-D gel systems, indicating that these molecules are covalently sealed in a hairpin structure. Treatment of thymocyte DNA with agents such as S1 nuclease that eliminate hairpin structures resulted in a dramatic reduction in molecules that migrate off the diagonal. Furthermore, a modified LM-PCR that involved the use of agents that disrupt hairpins prior to ligation, demonstrated that the coding ends terminate exactly at the sequence adjacent to the RSS of the germline element (132). This finding provided critical support for the hypothesis that cleavage occurs directly between the 5' end of the heptamer and the coding element.

Controversy quickly arose as to whether the isolated hairpin structures in SCID represented true recombination intermediates. It was argued that these
hairpin structures could be a consequence of the defective DSB repair pathway characteristic of SCID mice, and accounted for the inability to detect hairpins in normal mice (133). Others postulated that the failure to detect these hairpins in normal lymphocyte progenitor cells was a consequence of rapid resolution of these structures into coding junctions (130, 134).

Several lines of evidence suggest that coding end hairpin structures are true recombination intermediates. First, hairpinned coding end intermediates provide an attractive explanation for the existence of P nucleotides in coding junctions (88, 130, 135), and their absence in signal junctions. Secondly, artificial hairpin recombination substrates introduced into both SCID and wild type cell lines are readily joined and these junctions exhibit P insertions (42). Third, unmodified hairpin coding ends and full length blunt RSS DSBs from the endogenous κ locus are observed in temperature sensitive A-MuLV lines shifted to recombination favourable conditions (131). Interestingly, open processed coding ends were detected in these cells. None of the open ends isolated terminated exactly at the end of the coding element, indicating that this species of DNA end are products rather than precursors to the hairpin coding ends observed. The observed coding end hairpin structures and RSS DSBs may still represent an aberrant diversion of recombination given that they arise in transformed cells at non-physiological temperatures. However, these same conditions give rise to coding junctions that are indistinguishable from those in normal lymphoid cells. Moreover, these coding joins accumulate with similar kinetics to signal ends, indicating that the coding end resolution is rapid, and occurs independently of signal ends joining (131). Interestingly, open processed coding ends, but not covalently sealed coding ends, are isolated from the TCRα locus of normal thymocytes (136). This finding may imply that coding end hairpin structures are rapidly processed and do not accumulate to a significant
degree under physiological conditions. However, this data is also consistent with the existence of an alternative recombination pathway that does not involve a coding end hairpin intermediate.

**In vitro Systems**

For many years, the biochemical mechanism underlying V(D)J recombination could only be inferred from the analysis of its end products made in cells. It was apparent that the expression of the RAG proteins was critical in conferring V(D)J recombination activity. However, it was not known whether these gene products were direct participants in the recombination process, or whether they served to activate the "recombinase" proteins. The RAG proteins were implicated in the cleavage reaction by the observation that both proteins are required for DSB formation *in vivo* (60). Further evidence implicating the RAG proteins as direct participants in recombination was derived from the identification of mutant RAG-1 proteins that are particularly sensitive to the nature of coding sequences flanking the RSS (123, 124).

Experiments directly addressing the issue were stifled by the inability to attain soluble full-length RAG proteins. Mutational analysis of the RAG genes identified active protein cores that are sufficient to confer V(D)J recombination activity in transfected cells, and are more soluble in nature (65, 66, 68, 69). The ability to express soluble RAG proteins coupled to the identification of V(D)J recombination intermediates served as catalyst for the development of an *in vitro* RSS-specific cleavage system.

Initially, nuclear extracts supplemented with purified truncated RAG-1 protein generated identical DSBs and hairpins observed *in vivo* (137). This activity correlated with nuclear extract fractions containing RAG-2 protein. Further refinements of this technique demonstrated that only purified core RAG-1 and core RAG-2 proteins are required for cleavage and hairpin formation to
**Figure 1-9:** V(D)J Recombination Cleavage Reaction. A. Schematic diagram depicting nicking of an RSS oligo substrate, and subsequent hairpin formation of the coding sequence. B. Pre-nicked substrate with the phosphate group removed three nt from the 5' end of the heptamer still forms a hairpin at the appropriate site adjacent to the heptamer. C. Single stranded RSS oligo substrate can still form hairpins, however complete hairpin formation is inaccurate, and limited to the borders of the heptamer.
occur (125). Moreover, the quality of the enzymatic reaction is dependent upon the divalent cation used. In the presence of Mn\(^{+2}\), cleavage and hairpin formation readily occurs at a single RSS (125, 137), however two RSS of complementary spacer length are required for cleavage in the presence of Mg\(^{+2}\) (138, 139).

The process of cleavage occurs in two separate stages (125): a nick is introduced immediately 5' of the heptamer's position one residue, generating a 3' hydroxyl group on the adjacent coding sequence and a 5' phosphate on the 5' end of the heptamer (see Figure 1-9A). The process of nicking is rapid, occurring within two minutes and saturates by 20 minutes after the reaction is started. The second step is the conversion of the nick into a hairpin by a nucleophilic attack of the opposite strand, subsequently generating a blunt 5' phosphorylated signal end and a hairpinned coding end. Hairpin formation is slower than the nicking process, requiring at least eight minutes before detection, and saturates by six hours.

Cleavage does not require ATP or any other high-energy cofactor, suggesting that either energy derived from the initial nick or from the disruption of the phosphodiester bond on the opposing strand is sufficient for hairpin production (125, 137). This issue was addressed using pre-nicked substrates that involve nicks adjacent to the upper strand 5' end of the heptamer, i.e. an introduction of a 3' hydroxyl group on the flanking coding sequence (125). These pre-nicked substrates are readily converted into hairpins. Thus, energy derived from the disruption of the phosphodiester bond during the initial nicking event is not required for hairpin formation. Interestingly, prenicked substrates in which the 3' hydroxyl group is removed from the heptamer by up to three nucleotides on the upper strand (i.e. 5' to 3'), can still form hairpins readily (Figure 1-9B). Moreover, nucleophilic attack always occurs at the
phosphodiester bond between the heptamer and coding element on the bottom strand (3' to 5') (126, 127). Prenicked substrates containing a 3' hydroxyl group at position one on the bottom strand of the oligo substrate heptamer, are not converted into hairpins (125). The lack of hairpin formation on these prenicked substrates likely reflects the manner in which the RSSs are bound by the RAG proteins.

Both the initial nicking and hairpin formation require a RSS. Prenicked substrates with random sequences in lieu of a RSS are not covalently sealed. Furthermore, mutations introduced at position one of the heptamer on the upper strand of the RSS results in inaccurate cleavage, and prevents hairpin formation. The incidence of nicking is also reduced two-fold when sequences of position one and two of the heptamer are altered from consensus (126).

**Synapsis**

Efficient V(D)J recombination requires rearrangement to occur between coding elements flanked by RSSs of differing length, this phenomenon is referred to as the 12/23 rule (30, 32, 33, 50, 140, 141). Moreover, the products of recombination clearly indicate that the recombining elements pass through a phase in which they are synapsed. Several lines of indirect evidence suggest that synapsis of the recombining elements occurs prior to cleavage. First, the frequency of recombination is severely reduced when the target RSSs in artificial substrates are separated by short stretches of DNA (≤100 bp) which increase the amount of energy required for DNA bending (33, 50, 142). RSS cleavage in a cell free system is also substantially inhibited by short intersignal distances (138), indicating that synapsis must precede cleavage. Second, the frequency of open and shut junctions which theoretically only require recognition and cleavage of one RSS, is enhanced when a pair of complementary RSSs are present (50, 143). Third, mutations at one RSS appear to significantly interfere with cleavage at the
complementary consensus RSS in both in vitro RSS cleavage assays and in vivo recombination substrate systems (54, 138, 139).

It appears likely that this synaptic complex involves additional proteins. First, introduction of RAG genes into fibroblast lines are insufficient to support rearrangement of the endogenous TCR and Ig loci. Secondly, recombinant core RAG proteins are unable to initiate cleavage of nuclear DNA and purified genomic DNA (144). However, this activity can occur if nuclear extracts derived from recombinationally active lymphoid cells are added (144). The auxiliary factors appear to be lymphoid specific as HeLa cell extracts are unable to complement the RAG proteins in this assay. The nature of these proteins is unknown as yet.

Synapsis however is not mandatory for cleavage to occur. Open and shut junctions occur in single RSS substrates at similar frequencies to substrates containing pairs of complementing and matching RSS (50). Moreover, low levels of cleavage are consistently observed in single RSS substrates or substrates containing two RSSs of the same spacer length (54).

**STAGE III RESOLUTION**

The biochemical events governing the resolution of the recombination intermediates are the least characterized of the three phases of recombination. Our understanding of this phase of recombination is inferred from the quality of the end products, and from the identification of various mutants unable to complete the process.

TdT is involved in the resolution process in immature lymphocytes of adult mice, but is not essential for recombination. Mice deficient in TdT do not have a defective lymphoid development, and do not appear compromised in their ability to mount an immune response (70, 71, 80).
The profound deficiency of lymphoid development in SCID mice is a consequence of inability to resolve coding ends ((38), see SCID below). Subsequently, the scid defect was discovered to be ubiquitous in nature in that it also compromised DSB repair (3-5, 145). The realization that the scid mutation affects both DSB repair and V(D)J recombination catalyzed the search for other mutations with similar properties. Several mutants comprising three other DSB repair complementation groups were isolated (146). The nature of these mutants is discussed in detail below (see SCID). Interestingly, all three of these complementation groups exhibit a more severe defect in V(D)J recombination than the SCID mutation, as signal junctions are also impaired.

B CELL DEVELOPMENT & Ig REARRANGEMENT

ONTOGENY OF B CELL DEVELOPMENT

Hematopoietic stem cells (HSC) are defined by their capacity to permanently repopulate multiple hematopoietic lineages in lethally irradiated mice, and by their ability to self-renew. The fetal liver serves as the major site of hematopoiesis during embryonic development. B lineage committed precursors are detected as early as day 12 of gestation in mice (147). By day 16 of gestation, the site of hematopoiesis shifts from the liver to the BM, where it remains throughout adult life (148). B cell development in the BM is initiated in the regions next to the bone. Progression through the B lineage is accompanied by a migration of the progenitor B cell towards the central region of the BM (as reviewed in (149)).
Ig REARRANGEMENT AND B CELL DEVELOPMENT

Commitment to the B cell lineage is measured by the capacity of a cell to differentiate and proliferate in response to stromal cells and/or cytokines, and ultimately by the cell's Ig gene rearrangement status (as reviewed in (150)). B cell development can be subdivided into four general phases based on the Ig rearrangement status: pro-B, pre-B, immature B and mature B, and are discussed in detail below (see Figure 1-10 and Table 1-1). These criteria, coupled with the establishment of in vitro culture systems which promote maturation of B lineage committed cells (76, 77, 151, 152), permitted the identification of cell surface and intracellular proteins whose expression is associated with the various stages of B cell development (76, 77, 106, 153, 154).

• Pro-B Cells

Pro-B cells are the earliest identifiable cells committed to the B cell lineage (76-78, 93, 106). The IgH loci of these cells are characteristically in either a germline configuration or DJH rearranged configuration (76, 78, 93). Pro-B cells are unresponsive to IL-7 in the absence of contact with stromal cells, and consequently fail to proliferate and undergo further maturation (76). The requirement for stromal cell contact can be overcome by MGF and IL-11, indicating that stromal cells do not provide an intrinsic contact dependent signal for further differentiation (150).

Like all cells committed to the B lineage, pro-B cells express the B220 isoform of CD45 (76, 77, 106, 147). Pro-B cells are further defined by the expression of CD43 and ckit, but lack BP-1 and heat stable antigen (HSA) which are characteristic of more mature stages of development (76, 77). The low level RAG and TdT gene expression coupled to the few detectable DJH rearrangements in these cells is consistent with commitment to the B lineage (77, 78). However, the B220+CD43+BPl−HSA− cells are heterogeneous in nature; only
a fraction of these cells can proliferate and differentiate in response to contact with a stromal cell layer (76). Differential expression of the hematopoietic marker AA4.1 has permitted identification of B lineage committed cells within this population (106). Only AA4.1+ cells can proliferate and differentiate on stromal cells. This population can be further subdivided by the expression of CD4, a marker commonly associated with the T cell lineage but is also expressed at low levels on lymphoid progenitor cells (155-158). The B220+AA4.1+CD4lo cells are referred to as Fraction A1 in the Hardy scheme of B cell development (106), and are designated as pro-BI cells here (see Table 1-1). High levels of germline IgH transcription, and low, but detectable RAG gene transcription in these cells provides evidence of commitment to the B lineage, and indicates that accessibility of the IgH locus precedes functional recombinase activity in vivo (106). Moreover, detection of low levels of B29 (Igβ) gene transcription is also consistent with commitment to the B lineage.

Further progression in the B lineage is associated with the loss of CD4 expression (106). These cells are referred to as Hardy's Fraction A2, and as pro-BII cells here. Pro-BII cells exhibit increased expression of the RAG and B29 genes (106). Moreover, other B lineage and lymphoid specific proteins are now expressed, including λ5, a component of the surrogate light chain (see below), and mb-1 (Igα) and TdT (77, 106). This fraction of cells likely holds the DJH and the rare VDJH rearrangements that are detectable by single cell PCR analysis of the heterogeneous B220+CD43+ckit+BP1-HSA- (77, 78, 93).

• Pre-B Cells

Pre-B cells were first identified by their expression of cytoplasmic μ (cμ) heavy chains in the absence of surface Ig expression (159). Consequently these cells contain at least one VDJH rearrangement. This phase of B cell development is also heterogeneous, and can be further subdivided into two sub-stages: late
Figure 1-10: B cell development adapted from (10).
### TABLE 1-1
Summary of Murine B cell Development

<table>
<thead>
<tr>
<th></th>
<th>Pro-B</th>
<th>Pre-B</th>
<th>Immature B</th>
<th>Mature B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pro-BI</td>
<td>Pro-BII</td>
<td>Pre-BI</td>
<td>Pre-BII</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B220</td>
<td>+</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>AA4.1</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td>CD41&lt;sup&gt;+&lt;/sup&gt;</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CD43</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>ckit</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>HSA</td>
<td>-</td>
<td>+</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>BP-1</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CD25</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
</tr>
<tr>
<td>cμ</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>IgM</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IgD</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>mb-1 (Igα)</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>B29 (Igβ)</td>
<td>-</td>
<td>+</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>Il-7R</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>RAG-1/2</td>
<td>+/-</td>
<td>+</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>TdT</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Vpre-B</td>
<td>-</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>λ5</td>
<td>-</td>
<td>+++</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Bcl-2</td>
<td>+++</td>
<td>++</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

B

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DJH</td>
<td>+/-</td>
<td>+++</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>VDJH</td>
<td>-</td>
<td>+++</td>
<td>+++</td>
<td>+/+</td>
</tr>
<tr>
<td>VJK</td>
<td>-</td>
<td>+/-</td>
<td>++</td>
<td>+++</td>
</tr>
</tbody>
</table>

A. Gene Expression During B cell Development
B. Rearrangement activity in at a particular Ig locus
+/- indicates protein expression and/or activity occasionally is observed at the indicated stage of development. See text for details.
pro-BII/pre-BI, which are VDJH rearranged, but lack cμ protein expression, and the cμ+ pre-BII cells (160).

Late pro-BII/pre-BI cells are phenotypically distinguished from pro-BII cells by a gain in surface heat stable antigen (HSA) expression, and are also referred to as Fraction B under the Hardy scheme of B lineage development (76). Cells within this population exhibit a moderate dependence upon stromal cells for viability, and can proliferate in response to IL-7 alone. Increased expression of RAG, TdT, and the surrogate light chain genes (λ5 and VpreB) is characteristic of these cells (77). Moreover, these cells exist primarily in DJH/DJH and VDJH/DJH configuration (93), which consequently explains the lack of detectable germline transcription of the IgH locus in these cells (106).

Pre-BI cells are marked by the acquisition of surface BP-1 expression (76, 77), and are also referred to as Hardy’s Fraction C. These cells can undergo a robust proliferation in response to IL-7 that is independent of stromal cell contact (76). Pre-BI cells exhibit extensive VDJH rearrangement; the majority of which are nonproductive on both alleles which accounts for the lack of detectable cμ protein (78, 93, 154, 161). However, the surrogate light chain genes are expressed at high levels similar to pro-BII cells (77). Moreover, surrogate light chain proteins are observed in association with a 130kDa glycoprotein (gp130), as well as other glycoproteins ranging from 35 to 65kDa, in both the cytoplasm and on the surface of pre-BI cells (160, 162) (see below). RAG, TdT, mb-1 and B29 gene expression is comparable to pro-BII cells (77).

The transition from pre-BI cells into large pre-BII cells is marked not only by the expression of cμ protein, but also by an increased expression in HSA, and a concomitant decrease in ckit and CD43 expression (76, 153, 161). These cells are also called Fraction C', and have an enhanced proliferative capacity when transferred into SCID mice (76). The pre-B cell receptor (pBCR) is exclusively
expressed on the cell surface of these large pre-BII cells (see below), of which 60-70% of these cells are in cell cycle (i.e. S, G2 or M) (92, 163, 164). Single cell PCR analysis has revealed that the vast majority of these cells contain one in frame VDJH rearrangement, whereas the IgL loci remain predominantly in germline configuration (78, 93, 94). These cells are further characterized by a lack of detectable RAG2 protein, as well as termination of TdT and RAG gene transcription (92). Increased expression of CD25, the α chain of the IL-2 receptor marks a further differentiation within the large pre-BII cells (154). The majority (~60-70%) of these large CD25+ cells are also actively cycling; however, the pBCR is no longer expressed on the surface (92, 163, 164). The large CD25+ cells are characterized by an increase in RAG gene, as well as germline κ locus transcription (92). V(D)J recombination is likely not occurring in these cells as evidenced by the predominant germline configuration of the κ locus (94), and the lack of detectable RAG2 protein in these cells (92).

The small pre-BII cells comprise 75% of the pre-BII population (161). These cells are quiescent and no longer express surrogate light chain molecules (77, 163, 164). Moreover, these cells are characterized by increased RAG gene transcription, and active light chain gene rearrangement (77, 92, 94).

• Immature B

Immature B cells express IgM on the surface, but no IgD. These cells are unable to proliferate and mature into Ig-secreting plasma cells in response to Ag stimulation (as reviewed in (165)). Moreover, these cells are believed to require stromal cell dependent signals in order to proliferate in response to LPS. The RAG genes are no longer transcribed in immature B cells (77, 92), however both RAG proteins are still expressed (92). The continued expression of the RAG2 protein is consistent with findings that suggest immature cells undergo secondary rearrangements in the IgL loci (92, 166-168).
**Mature B**

Mature B cells are distinguished from immature cells by the increased expression of B220 and the coexpression of IgD (76). The RAG genes are not expressed in these cells, and RAG2 protein is undetectable (92). Mature B cells are further characterized by L-selectin lymph node homing receptors, complement receptors one and two, CD22, CD23 and CD40 expression (as reviewed in (165)). Under proper antigenic stimulation, these cells can be induced to differentiate into Ab secreting plasma cells and memory B cells.

**ALLELIC EXCLUSION**

Membrane Ig are essential for the regulation of both developing and mature B cells. The \( \mu \) chain controls pro-B to pre-B cell transition (169-172), allelic exclusion (173), and deletion of lymphocytes that are autoreactive.

B cell development is characterized by a relatively ordered rearrangement of the Ig loci variable region genes (174). The first stage of gene rearrangement is the juxtaposition of a DH element to a JH element. The resulting DJH junction then serves as a substrate for VH joining. In frame VDJH rearrangements result in the expression of the \( \mu \) chain. Light chain rearrangement generally follows (175-178). Production of a functional light chain permits surface IgM expression and further differentiation.

Mature B cells express only one pair of productively rearranged IgH and IgL chain genes, despite having the capacity to rearrange both IgH or all four IgL chain alleles. Allelic exclusion is the termed used to describe this phenomenon, while isotypic exclusion refers to the restricted expression of only one productively rearranged light chain isotype.

Several models were suggested to account for the allelic exclusion phenomenon. The stochastic model proposed that allelic exclusion is a
consequence of the low probability of producing two productively rearranged IgH alleles in a single cell (113). A second hypothesis suggested that the concurrent expression of two IgH chains is toxic for the cell (112). Both models require that Ig rearrangement occurs independently at alleles. A third hypothesis proposed that expression of a μ chain terminates further VH to DJH joining of the IgH locus (103). This hypothesis was derived from analysis of Ig loci configurations in transformed B cell precursors. Mice transgenic for the membrane form of the μ chain are allelically excluded while those bearing only the secreted form of the transgene are not (reviewed in (165)). Secondly, B cells derived from mice heterozygous for a disrupted μ chain membrane exon are allelically included (i.e. productive IgH rearrangements are found on both alleles), suggesting that membrane bound μ chains elicit a signal that prevents further VH to DJH joining (165). Mice engineered to express two distinct μ chains exhibit a normal sized B cell compartment in which most cells stable express both μ chains (179). Moreover, inspection of pre-BI cells in normal mice indicates that allelic exclusion is already in place upon expression of μ chains (93).

Collectively, these findings discredit the toxicity model described above.

The mechanism underlying this μ chain mediated allelic exclusion was unknown until the discovery of the surrogate light chain gene products. Association of these proteins with the μ chain and Igα/β permits surface expression of the μ chain in the absence of IgL chain.

• The surrogate light chain

The surrogate light chain in mice consists of two protein molecules, λ5 and VpreB. The genes encoding the λ5 and VpreB proteins are located on chromosome 16, along with λ light chain locus (as reviewed in (180)). These genes are selectively expressed in precursor B lineage cells (discussed above).
The 18 kDa λ5 protein bears AA similarity to Jλ-Cλ. It covalently associates with the μ chain by a disulfide bridge (181-183). The 22 kDa VpreB protein is similar to Vλ, and associates with the μ chain by noncovalent interactions (181, 182). There are two isoforms of VpreB, VpreB1 and VpreB2, which differ by only four AA. These proteins are coexpressed in both pro-B and pre-B cells (184). It is unknown whether the two VpreB isoforms have different signaling capacity and/or ligands.

The λ5 and VpreB proteins associate with one another to form a light chain structure referred to as the surrogate light chain. The surrogate light chain can pair with three different structures to form Ig-like complexes on the surface of precursor B cells. These complexes include the glycoproteins (gp)130/gp35-65 (115, 164), the Dμ protein (183, 185) and the μ heavy chain (reviewed in (160)). All of these structures require Igα/β heterodimer for surface expression.

Surrogate light chain associates with non-μ chain gp130/gp35-65 are predominantly found on the surface of pro-B cells (164). This pro-B cell receptor is present on progenitor B lineage cells derived from RAG-2 deficient mice (59, 163). Initially, this receptor was hypothesized to signal both differentiation and initiation of V(D)J recombination (115). However, this role is unlikely given that mice deficient in λ5 expression have normal numbers of pro-B and pre-BI cells (186). Moreover, λ5-deficient B cells successfully undergo V(D)J recombination at the IgH, and even IgL loci (albeit at a much reduced frequency) (78, 163, 186, 187). Consequently, the functional significance of this receptor in B cell development remains unknown.

All but one DH element have promoters located upstream of the coding element (185). DJH rearrangements in reading frame (RF) 2 place the upstream promoter in frame with the JH element, and result in a truncated μ chain called Dμ (185). The Dμ receptor associates with the surrogate light chain and the Igα/β
heterodimer and is expressed only on the surface of pro-B and pre-BI cells (164, 188). Dμ surface expression appears to require the association of a fifth protein as VDJH/VDJH rearranged pre-B cell lines can not express Dμ on the surface (183). This protein is postulated to be derived from the germline transcripts of VH elements.

Selection against cells bearing DJH rearrangements in RF2 is believed to occur at this phase of development because of the absence of RF2 DJH rearrangements in subsequent stages of development (see below) (187).

Rearrangements involving DH elements in RF2 are rarely found in either VDJH or DJH joins of pre-B and mature B cells (93, 189, 190). Most productive VDJH rearrangement are in RF1. In the absence of TdT expression, RF1 is preferentially used because of homology mediated resolution (26, 82-84, 191). However, homology mediated resolution cannot account for the RF1 bias in the adult repertoire. Junctions in which the DH elements are in RF3 have a high incidence of stop codons (24), thus accounting for their absence in mature B cells. Indeed, analysis of nonproductively rearranged IgH alleles yielded equal RF1 and RF3 usage (192).

The counter selection of Dμ receptor bearing cells is believed to be a consequence of signals transduced through the Dμ receptor. This signal is thought to prevent VH to DJH joining. The finding that RF2 rearrangements are absent or rare even in nonproductively rearranged VDJH structures present on the second allele of normal cells provided evidence for this notion (93). More evidence comes from cells unable to express the Dμ receptor (μ membrane deficient (μMT) (190), λ5−/− (187) mice) or to transduce signals through this receptor (syk−/− (193), Igβ−/− (194)) have equal RF usage in their Ig rearrangements.
The pre B cell receptor (pBCR) consists of a fully rearranged μ chain protein, the surrogate light chain and the Igα/β heterodimer (181, 182, 195). pBCR expression appears to be confined to the large pre-BII stage of development where rapid proliferation is occurring (164). The large pre-BII cells in normal mice contain only one productive VDJH rearrangement, indicating that expression of the pBCR, like Dμ, prevents endogenous VH to DJH rearrangement (93). The pBCR's role in allelic exclusion is highlighted in heterozygous μMT mice where many of the peripheral B lineage cells contain productive rearrangements on their mutated allele (196).

The pBCR plays a critical role in B cell development; mice unable to express the pBCR, homozygous μMT mice, λ5- and Igβ-deficient mice, exhibit a block in the transition from pre-BI to pre-BII cells resulting in either a drastically reduced or complete absence of B cells (176, 186, 197). A comparable block in B cell differentiation is also observed in SCID (6), and RAG-deficient mice (58, 59). The common arrest in B lineage differentiation illustrates the importance Ig expression plays in directing B cell development.

The Ordered Rearrangement Hypothesis postulated that the pBCR also acts to direct rearrangement of the IgL loci (175, 177, 198). However, κ rearrangement occurs in the absence of heavy chain expression (104, 196, 199). Moreover, elegant studies of wild type, μMT, JH locus- and λ5-deficient mice revealed that κ rearrangement can occur independently of IgH rearrangement and surface pBCR expression (78, 93). These studies suggest that the pBCR does not control the process of light chain gene rearrangement, but rather is essential in producing a large number of cells that normally become competent to undergo light chain rearrangement.
THE SCID MOUSE

DISCOVERY AND PHENOTYPE

SCID mice were first identified in 1980 by the fortuitous observation that littermates among the C.B-17 inbred strain of mouse (BALB/c·C57BL/Ka-IgH-1b/ICR (N17/F34)) lacked detectable serum Ig (6). Further investigation demonstrated that these mice are devoid of both B and T cell mediated functions, as they are unable to respond to T-independent antigens, mitogenic stimuli and allogenic skin grafts (6, 200, 201).

Selective breeding determined that this defect is encoded by an autosomal recessive gene, and permitted the establishment of a strain of mouse homozygous for the scid defect (SCID) (6). The scid mutation is linked to mahoganoid gene (md) which encodes a recessive coat colour located on the centromeric end of chromosome 16 (202).

• SCID and Lymphoid Development

SCID mice are leukopenic (6). Secondary lymphoid organs such as the spleen and lymph nodes are only 10% the size of normal and are virtually devoid of mature lymphocytes (6). These tissues are identified by their basic stromal structure and sinus patterns. The lymphoid follicles and germinal centers characteristic of lymph nodes and spleen are replaced by a nodular structure consisting of stromal cells and varying proportions of macrophages and granulocytes (1).

The SCID thymus is also hypoplastic in nature, owing to a lack of the more mature thymocytes. T cell development in SCID mice is arrested at the CD4/CD8 double negative (DN) stage (6). Approximately 10% of SCID mice aged five to nine months develop thymomas (6, 203). Analysis of TCR rearrangements in these thymomas revealed gross deletions comparable to those
observed in A-MuLV transformed SCID B cells (39) (See The SCID Defect and V(D)J Recombination below). Primary SCID T lineage cells characteristically undergo germline transcription of their β, γ and δ loci (204), indicating that the progenitor cells become competent to attempt V(D)J recombination.

SCID BM is histologically normal. However, SCID BM does exhibit a profound block in B cell development at the pro-B cell stage (6). SCID mice contain normal numbers of pro-B cells which express RAG proteins, and undergo germline transcription of the IgH loci (204-206). Indeed SCID mice have comparable numbers of A-MuLV transformable cells as wild type mice, revealing that the target cell of the A-MuLV is the pro-B cell (145). However, neither transcripts from fully rearranged IgH loci nor μ+ pre-B cells are detectable in these mice (6, 204, 206, 207). Moreover, histological and functional analysis revealed that differentiation and function of the myeloid and natural killer (NK) cell lineages are unaffected by the scid defect (6, 200, 201). Collectively, these data indicate that the scid defect specifically impairs lymphoid development.

The block in early lymphoid cell maturation suggested that the scid defect compromises lymphoid development rather than the functional capacity of lymphocytes. Reciprocal reconstitution studies between BALB/c and SCID mice demonstrated that the SCID mice are fully capable of supporting lymphoid development (6, 208), whereas SCID progenitor lymphoid cells fail to differentiate in normal mice. Thus, the scid defect is a property inherent to the lymphoid lineage, and not due to defects in the SCID microenvironment.

• SCID Mice Are Leaky

The SCID defect is not complete; 2-25% of SCID mice develop evidence that functional B and T cells had to have been generated (7, 209, 210). This leaky phenotype was first identified by the detection of serum Ig, and the restoration of
some lymphoid effector functions such as allograft rejection and mitogen responsiveness (7). The incidence of leakiness correlates with both age and environment, as older mice and those mice raised in non-specific pathogen free conditions have a higher incidence of detectable Ig (7). The leaky phenotype is not inheritable, and thus does not appear to be a consequence of a germline genetic reversion event (7). Moreover, BM cells from leaky SCID mice transferred into irradiated wild type mice fail to generate Ig producing cells (7).

The concentration of serum Ig varies greatly among different SCID mice (7, 211). Leaky SCID mice generally exhibit isotype switching, however all the isotypes are rarely present in a given mouse. The serum Ig present in leaky SCID mice is generated by only a few distinct B cell clones. Isoelectric-focusing (IEF) patterns of serum κ light chains derived from serum Ig positive SCID mice typically contain less than ten unique κ light chains (7, 211, 212). Moreover, the IEF patterns vary considerably between SCID mice. However, the IEF pattern for an individual SCID mouse is generally stable over time, suggesting that the generation of new functional B cells is minimal, and that existing clones are relatively long lived (7, 211).

Serum Ig positive SCID mice generally contain a limited number of T cells in both the periphery and thymus (7, 209, 210). The incidence of T cell leakiness is measured by the ability of these mice to reject allografts, and to respond to mitogenic stimuli such as concanavalin A (Con A) (7, 209). Many leaky SCID mice are capable of only partially rejecting the allograft, suggesting that the number of distinct functional T cells produced is greatly limited. Furthermore, FACS analysis of leaky SCID mice revealed only a limited number of detectable CD3 positive T cells in the thymus and/or peritoneal cavity and spleen (209, 210). Some of the leaky SCID mice contained only CD4 or CD8 single positive (SP) T cells. Nonetheless, the oligoclonality of lymphoid development in leaky SCID
mice suggests that these lymphoid cells arise by a rare somatic event in a cell that possesses limited self-renewing capacity (7, 209-212). Interestingly, the incidence of T cell lymphoma is almost four-fold higher in leaky SCID mice as compared to non-leaky SCID mice (7).

• Conditions Affecting Lymphoid Development in SCID Mice

Several experimental manipulations can circumvent the profound block in SCID lymphoid development: (1) transgenic expression of rearranged Ag receptor transgenes (2) adoptive transfer of wild type BM or thymocytes, (3) abrogation of NK and monocyte cell activity (4) the introduction of the anti-apoptotic bcl-2 transgene, and (5) irradiation of newborns with sublethal levels of γ-radiation.

Antigen Receptor Transgenic Mice

Introduction of fully rearranged Ig or TCR transgenes into SCID mice induces variable levels of lymphoid development (170, 172, 213-215). The extent to which a transgene promotes lymphoid differentiation correlates with its capacity to inhibit endogenous rearrangement (172). The presence of a single transgene alone i.e. IgH or κ, does not result in enhanced leakiness in these SCID mice.

Adoptive Transfer of Wild Type BM or Thymocytes

TCR- SCID thymocytes can be induced to differentiate into CD4+/CD8+ cells upon adoptive transfer of wild type T-depleted BM cells into non-leaky SCID mice (216). The CD8 isoform expressed on the SCID thymocytes is phenotypically distinct from that expressed by the donor wild type thymocytes. Thus, the SCID-derived thymocytes enter into the CD4/8 differentiation pathway themselves, rather than passively acquiring the CD4/8 markers from wild type cells. These data suggest that TCR+ cells provide the necessary signals within the
thymic microenvironment to promote entry of DN TCR− cells into the CD4/8 differentiation pathway in the absence of TCR expression.

Transfer of neonatal thymocytes into non-leaky adult SCID mice results in high levels of host Ig secretion in all SCID recipients (217-219). The capacity to promote host Ig secretion in SCID mice is exclusively conferred by CD4+ thymocytes derived from mice less than three weeks of age (218); adult-derived thymocytes promote only minimal host Ig secretion (217). Moreover, only donor neonatal cells that share MHC-identity are capable of inducing high levels of host Ig secretion, suggesting that a mature T helper cell is responsible for the induction of Ig secretion (218, 219).

Serum Ig isotype levels vary among different recipient SCID mice (217), and these mice fail to mount specific humoral immune responses (219). Interestingly, transfer of neonatal T cells into older non-leaky mice results in substantially higher production of host serum Ig than younger recipients indicating that SCID mice amass functional B cells over time (217). These results suggest that transfer of neonatal T cells does not provide factors necessary to override the scid defect, but rather, these T cells appear to rescue the few SCID B cells that arise (217-219). The developmental limitation of donor cells coupled to the requirement of syngenic donors to support host Ig secretion suggests that the donor cells are likely auto-reactive T cells that have yet to be anergized (218).

Abrogation of NK and Monocyte Activity

Repeated administration anti-asialo GM-1, an antiserum directed against NK cells and macrophages, induces serum Ig production in 40% of non-leaky SCID mice (220). Similar treatment of SCID mice engrafted with human thymus also results in 38% of these mice developing serum Ig. Analysis of the human grafts in these mice revealed only murine SP TCR+ thymocytes to be present. Moreover, detection of murine T cells in the thymus and spleen of these
engrafted, anti-asialo GM-1 treated mice, is correlated with the presence of serum Ig. These findings suggest that NK cells and/or macrophages may act to suppress lymphopoiesis in SCID mice by removing autologous lymphoid precursors (220).

**SCID/Bcl-2 Transgenic Mice**

Overexpression of the bcl-2 proto-oncogene under the Eμ promoter (EμPr) in SCID mice promotes B cell differentiation, whereas T cell development remains blocked at the DN stage (221). Near normal numbers of sIg- B lineage cells displaying other markers characteristic of mature Ig+ B cells accumulate in the SCID/(EμPr)bcl-2 mice. These SCID/(EμPr)bcl-2 B lineage cells exhibit extensive normal-sized DJH gene rearrangement products, and VDJH and κ rearrangement is frequently detected. The predominant VDJH rearrangements in these mice are abnormal in size, indicating that the enhanced survival mediated by the bcl-2 transgene does not compensate for the scid defect (221). Interestingly, IgH protein is detectable in B220+sIg- splenocytes of another SCID/bcl-2 transgenic mouse (222). However, it is currently unclear whether the protein detected represents Dμ or complete μ heavy chain protein.

Overexpression of bcl-2 per se is not sufficient to promote B cell development. RAG deficient/bcl-2 mice accumulate increased numbers of pro-B cells, but lack cells of a more mature B cell phenotype (222). However, expression of an IgH transgene in the RAG-deficient background promotes the transition of pro-B cells into pre-B cells (169, 171), and some of these pre-B cells express markers associated with a more mature phenotype i.e. CD22 and CD40 (222). Furthermore, the extent of IgH mediated differentiation in the RAG deficient background is enhanced by the concomitant expression of bcl-2 which results in greater numbers of peripheral pre-B cells that bear additional surface markers of more differentiated B cells i.e. CD23 (222). Collectively, these data suggest that
in the context of deregulated \textit{bcl-2} expression, IgH chain expression is essential to promote maturation of pro-B cells \textit{in vivo} (222).

The failure of \((\text{E} \mu \text{Pr})\text{bcl-2}\) transgene to promote thymocyte development is not due to a lack of transgene expression within the T cell lineage, as SCID/(\text{E} \mu \text{Pr})\text{bcl-2} thymocytes are reported to express comparable levels to wild type mature T cells (221). Moreover, introduction of a TCR transgene into these SCID/(\text{E} \mu \text{Pr})\text{bcl-2} mice permits the development of normal numbers of CD4+/CD8+ T cells, even in the absence of selection. This finding suggests that progenitor T cells become responsive to \textit{bcl-2} protein only after surface expression of the TCR complex. However, thymocytes of RAG-1 deficient mice transgenic for \textit{bcl-2} under the \textit{lck} promoter (\textit{lckPr}) differentiate into the DP stage, albeit very inefficiently (223). The disparity in the ability of the \textit{bcl-2} transgenes to support T cell differentiation in these two strains of mice likely reflect the different stages at which the two mutated genes influence the gene rearrangement process. For example, SCID DN thymocytes express a high level of p53 protein, the tumour suppressor gene that plays a critical role in both cell cycle control and apoptosis, whereas p53 is undetectable in both RAG deficient and wild type DN thymocytes (207). Moreover, RAG deficient/SCID double mutant thymocytes fail to express p53 protein, demonstrating that the upregulation of p53 in SCID thymocytes results from an inability to resolve the RAG initiated DSBs (207). Secondly, while T cell development is arrested at the DN phase in both SCID and RAG deficient mice (6, 58, 59), parallel analysis of these mutants reveal that T cell development is not blocked at the identical stages (224). RAG deficient thymocytes are quiescent, and accumulate at a single defined arrest point that phenotypically resembles the TCR\(\beta\) selction check point of normal thymocytes (224). SCID thymocytes however, die at more heterogeneous, and generally earlier stages within the DN phase of development (224). Expression of
the cell death triggering Fas receptor is observed on the more mature SCID thymocytes, but is undetectable in RAG deficient thymocytes (224). Thus, the increased expression of p53 protein (207) and/or upregulation of the Fas receptor (224), may account in part for the early cell death observed in SCID thymocytes, and the inability of the bcl-2 transgene to support T cell differentiation in the context of the scid defect.

Sublethal γ-Radiation

Treatment of neonatal SCID mice with a sublethal dose of γ-radiation (100-150 cGy) or other agents that cause DSBs, overcomes the developmental arrest of thymocytes and promotes progression into the DP stage (225-227). This limited rescue of T cell differentiation is accompanied with polyclonal TCR β rearrangements that exhibit normal deletions, lack the long P insertions characteristic of TCR rearrangements derived from leaky SCID mice, and most are in frame (see The SCID Defect and V(D)J Recombination) (226). Sublethal irradiation also permits the "normal" resolution of γ and δ rearrangements (227, 228), suggesting that a general mechanism supporting recombination is induced upon irradiation (228, 229). However, this mechanism appears to be transient in nature as VJα rearrangements are rare (226, 228, 230). Subsequent administration of radiation at a stage where TCRα rearrangement is initiated in the irradiated SCID mice fails to rescue TCRα coding joint formation and consequently complete T cell differentiation (230). These findings imply that the presence of hairpin coding ends at the time of irradiation per se is not sufficient to rescue their resolution. However, DP T cells are highly sensitive to irradiation as evidenced by the disappearance of most of the DP SCID cells within 24 hours of the second dose of radiation (230).

Radiation-induced T cell differentiation into the DP thymocytes is not dependent upon completed V(D)J recombination. RAG-deficient mice subjected
to irradiation also exhibit progression into the DP stage (231, 232). Thus, sublethal irradiation induces a TCRβ-independent differentiation pathway into the CD4/8 lineage. Subsequent analysis has demonstrated that this radiation-induced differentiation is dependent upon signals transduced by the src family tyrosine kinase, lck (233).

**MENAGE à TROIS: SCID, V(D)J RECOMBINATION & DSB REPAIR**

Concomitant with the discovery and characterization of the murine scid mutation in the early 1980's, the Ku autoantigen was identified by sera derived from patients with sclerodermatopolymyositis overlap syndrome (reviewed in (234), see Ku). The Ku protein complex was later shown to bind ds DNA ends regardless of the DNA sequence, and serves as the DNA binding domain of the DNA dependent protein kinase (DNA-PK). By the mid-1980's, a series of radiosensitive hamster mutant cell lines were isolated, and subsequently found to be defective in DSB repair. These three independent areas of intensive research dramatically converged upon the discovery that the scid defect affects V(D)J recombination and DSB repair. This discovery served as a catalyst in the identification of the SCID protein, and stimulated a breakthrough in our understanding of the mechanism(s) governing recombination and consequently lymphoid development.

**THE SCID DEFECT AND V(D)J RECOMBINATION**

The block in SCID lymphoid development is a consequence of a defective V(D)J recombination system. SCID A-MuLV transformed cell lines, long term bone marrow cultures (LTBMC) and spontaneously arising T cell lymphomas characteristically exhibit abnormally large deletions in attempted
rarrangements to J elements (39, 47, 235-238). These deletions often result in the complete removal of the entire J locus, and thus provides an explanation for the lack of mature lymphoid cells in SCID mice (47, 235, 236). The resolution sites of these aberrant rearrangements are indiscriminate in nature, implying that these joins arose from random deletion rather than recombination events to cryptic RSSs (235, 237).

Some SCID A-MuLV lines undergo endogenous Igκ rearrangement despite a lack of functional IgH rearrangement (199, 239). This observation was among the first evidence to indicate that light chain rearrangement was not strictly dependent on the presence of functional IgH chains in mice. Many of these endogenous SCID κ rearrangements occurred by inversion. Interestingly 80% of these SCID inversions contained perfectly normal signal junctions, while the reciprocal coding joins exhibited large deletions (199, 240). This finding was the first indication that RSS recognition in SCID is normal, and further accentuated the asymmetry of resolution of DNA ends during recombination. Moreover, these observations suggest that the mechanisms governing signal and coding end resolution are distinct and can be uncoupled.

Analysis of exogenous recombination substrates introduced into SCID A-MuLV lines revealed that SCID signal junctions resolve at a lower but comparable frequency to wild type cell lines, whereas SCID coding joint formation is at least a 10^3-fold decreased (38). Almost 50% of the SCID signal joints contain deletions and/or N additions; an occurrence rarely observed in wild type cells (38). Coding joints were not recovered from either inversional or deletional plasmid recombination substrates introduced into SCID lines, however, nonstandard hybrid junctions were detected (38). Most of the SCID hybrid junctions isolated contain uncharacteristic deletions in both their signal and coding ends. Nonetheless, the detection of hybrid joints in SCID cells
suggested that lack of coding joint formation is not a result of lost or destroyed coding ends in SCID cells.

Chromosomally integrated recombination substrates introduced into SCID A-MuLV lines also exhibit defective coding joint resolution (44, 239). The majority of the coding joints isolated involved large deletions, however, a few normal coding joints were detected (44, 235, 239). Aberrant secondary rearrangement events within the integrated substrates demonstrate that a genetic reversion event of the SCID mutation was not responsible for the observed normal coding junctions. Rather, this finding suggests that these junctions arose from a low frequency of normal recombinase activity (239).

A few studies suggest that some cells undergo a somatic reversion event. Some B and T cells derived from leaky mice have normal coding joint rearrangements on their non-expressed alleles (45, 241). Furthermore, an A-MuLV clone derived from the peritoneal cavity of a leaky SCID mouse expressed recombination activity indistinguishable from that of wild type cells (241). Alternatively, these observations may result from the activation of auxiliary repair enzymes that can compensate for the scid defect.

Analysis of Ag receptor genes derived from lymphoid cells in leaky SCID mice revealed that many of these junctions contain excessively long P insertions (47, 242). Little or no N addition was found in these rearrangements, despite normal TdT expression in SCID mice (242). Interestingly, about half of the rearrangements lacking N addition appeared to join by homology (242). It was unclear to what extent these features reflected Ag selection, and/or the recombination process in context of the scid defect. Several studies, including the contents of this thesis, have since shown that the relative lack of N addition, and increased frequency of P addition and homology-mediated joining is
characteristic of rearrangements derived from non-leaky SCID lymphoid cells (46, 243-246).

**SCID MICE ARE DEFECTIVE IN DSB REPAIR**

The understanding of how the SCID protein is involved in the resolution of coding junctions was revolutionized by the discovery that this defect was not limited to the lymphoid lineage. Non-lymphoid cells were shown to be hypersensitive to \( \gamma \)-radiation and radiomimetic drugs that induce DNA DSBs, but are resistant to UV irradiation and drugs which induce other types of DNA lesions, such as crosslinking or interstrand pyrimidine dimerization (2-5). Micronuclei accumulation as well as pulse field gel electrophoresis determined that the sensitivity to ionizing radiation was a consequence of a failure to repair the DSBs generated. Moreover, SCID cells are more vulnerable to the cytotoxic effects of DSB-inducing restriction enzymes introduced by electroporation (5). Interestingly, SCID cells subjected to \( \gamma \)-radiation exhibit an unusually high frequency of chromatid exchanges (247), this finding suggests that homologous DNA repair pathways are unaffected by the *scid* mutation and/or are induced in the absence of functional SCID protein (28).

- **The Break Connecting DSB Repair and V(D)J Recombination**

The isolation of coding end hairpin intermediates in SCID thymocytes, but not wild type T lineage cells, provided a possible link between the two defective processes in SCID mice (130). These hairpin intermediates from SCID mice could have represented either true V(D)J recombination intermediates, or an aberration from the normal rearrangement process due to its defect in DSB repair. As discussed earlier under MECHANISM, several lines of indirect evidence support hairpin coding ends as a V(D)J recombination intermediate. Briefly, a coding end hairpin precursor in the recombination pathway would
account for the presence of P nucleotides found in the coding junctions of both wild type and SCID mice. Secondly, intact hairpin structures are detected in a wild type A-MuLV line that gives rise to normal endogenous coding junctions (131). Moreover, linear processed coding ends have been isolated from both the wild type A-MuLV line (131), as well as from wild type thymocytes (136). These findings coupled to the observation that coding junctions accumulate with similar kinetics as signal ends (131), argue that coding ends are rapidly opened, processed and resolved, and thereby accounting for the lack of detectable hairpin coding ends in wild type thymocytes. Finally, \textit{in vitro} recombination assays which are capable of performing the initial RSS recognition and cleavage reactions produce blunt signal ends and covalently sealed coding ends that are comparable to those detected \textit{in vivo} (125, 131, 137, 139).

Qualitative analysis revealed that SCID signal ends are characteristically blunt and 5' phosphorylated, while the coding end structures are perfect, involving no alteration in the sequence of the original coding segment flanking the RSS (60, 88, 130, 132). These findings indicate that the \textit{scid} defect does not interfere with the initiation of the V(D)J recombination reaction, but rather suggests that the SCID protein is involved in the resolution of these structures. However, SCID cell lines resolve hairpin-ended, linearized plasmid DNA as well as wild type cells (42). The frequency and length of P insertions at the resulting junctions is comparable between the SCID and wild type cells. This result is consistent with other data in which linearized plasmid DNA ends containing a variety of terminal configurations are resolved at a frequency comparable to wild type cells (5, 38, 248, 249). Taken together, these findings indicate that the enzymatic components required for end-joining appear to be present in SCID mice (28). Moreover, these findings suggested that hairpin accumulation in SCID thymocytes was a consequence of an aberration from the normal V(D)J
joining process (42); a notion later dismissed upon the isolation of coding end hairpins in wild type cells (131).

The apparent discrepancy in the ability of SCID cells to resolve extrachromosomal and chromosomal lesions may not be as incongruous as it first appears. Other DSB repair pathways may operate independent of chromatin structure and/or a functional SCID gene product (5). For example, SCID cells can efficiently integrate plasmid DNA into chromosomal DNA suggesting that the nonhomologous DNA repair pathway is unaffected by the scid mutation (249). Alternatively, the SCID gene product may be involved in coordinating the resolution of recombination products with cell cycle and/or lymphoid differentiation (see below).

CHARACTERIZATION OF THE SCID DSB REPAIR MUTATION

In the mid-80's a genetic approach to defining components of DNA repair mechanisms in mammals involved the mutagenesis of Chinese hamster ovary (CHO) cell lines, and screening for sensitivity to ionizing radiation (IR) e.g. γ-rays and X-rays (250). IR involves high energy particles capable of releasing an electron from a atom or molecule, resulting in the generation of excited and ionized molecules (reviewed in (251)). This free radical formation results in a variety of DNA lesions in vivo that include both single strand and double strand DNA breaks, and chemically altered bases and sugars. Not surprisingly, the isolated CHO cell lines exhibit a spectrum of sensitivities to IR (234), likely reflecting both the pleiotropic nature of the damage induced by IR and the variety of different repair pathways required to repair the different types of lesions produced. Subsequent cell fusion studies identified at least eleven complementation groups (250). Each complementation group has been given an IR numerical designation. Five of these IR groups exhibit profound defects in
DSB repair (IR4-7), while two (IR1, IR3) involve defective single strand DNA break repair pathways.

Hypersensitivity to IR may result from either defective DNA repair pathway(s) or from an inability to elicit cell cycle arrest in response to DNA damage. For example, patients with ataxia telangiectasia (AT), a human autosomal recessive disorder characterized by a number of debilitating symptoms such as progressive neurological degeneration, immune deficiencies, growth retardation, are also hypersensitive to IR (as reviewed in (250, 252)). However, AT cells exhibit no "gross" defects in DNA repair. Rather, the IR sensitivity is believed to result from an inability to arrest cell cycle in response to DNA damage at both the G1-S and G2-M checkpoints. Indeed, AT cells exhibit abnormal p53 expression following DNA damage (253-257). (However, it is imperative to note that p53 checkpoint defect is unlikely to be responsible for pleiotropic features of AT, as p53 deficient mice are normal with the exception of increased incidence of lymphoma (258)). The scid defect appears to be distinct from that governing AT, as cell cycle arrest checkpoints are intact in SCID cells (207, 259-262). Moreover, complete restoration of radioresistance was achieved by fusing of SCID fibroblast cell line with a human AT fibroblast cell line (5).

The possible link between V(D)J recombination and DSB repair provided by analysis of SCID mice prompted the screening of these IR sensitive lines for the ability to support rearrangement. The competency of the various CHO lines to undergo V(D)J recombination was assessed by the concomitant transfection of RAG protein expression vectors with transient recombination substrates (146, 263-266). Interestingly, only mutants belonging to the complementation groups deficient in DSB repair (i.e. IR4 to 7, see below) were also defective in V(D)J recombination. Indeed, the AT-like IR8 mutant lines V-G8 and V-E5 which are characterized by normal single and double strand DNA break repair,
radioresistant DNA synthesis and chromosomal instability (250), exhibit normal signal and coding joint formation (263). Moreover, the IR10 BLM-2 cell line that exhibits slower single and double strand DNA repair (250), also permits normal V(D)J recombination (146). Consequently, IR sensitivity per se does not correlate with a V(D)J joining defect (263).

Interestingly, the severity to which V(D)J joining is affected differs among the four DSB repair complementation groups (summarized Table 1-2). Mutants belonging to IR groups four to six share a more profound defect in V(D)J recombination, as both signal and coding joints formation is dramatically reduced (146, 263, 266-268). Moreover, the rare signal junctions isolated from these mutants often display large deletions (146, 263). In contrast, the CHO V3 mutant belonging to IR7 is unable to resolve coding junctions, however both the frequency and overall quality of signal joint formation is only modestly affected in these cells (263, 264). V3 mutant cells share a comparable deficiency in V(D)J recombination to SCID cells; this cell line was subsequently demonstrated to belong to the same complementation group as SCID mice (264). The radiosensitive human malignant glioma M059J cell line, as well as equine SCID have also been identified as members of the IR7 group (269-271). However, equine SCID is unique among the IR7 members because both coding and signal joint formation are profoundly affected (270, 271). This apparent discrepancy will be discussed below.

No mutants defective in signal joint formation alone have been isolated to date. Nonetheless, the isolation of mutants with differing capacities in the resolution of coding and/or signal joints supports the hypothesis that coding and signal ends are resolved by different mechanisms (see V(D)J Recombination above). Moreover, these findings definitively demonstrate that V(D)J recombination involves the recruitment of ubiquitously expressed DSB repair
machinery. The existence of at least four complementation groups suggests that a multi-subunit complex and/or enzymatic pathway is utilized in the resolution of DSBs (234).

• Characterization of DSB Repair Mutants

Several strategies have been utilized in order to identify the proteins involved in DSB repair pathways. One such scheme involved interspecies complementation analysis. Using this technique human chromosomes 5q13-14, 2q33-35 and 8q11 were shown to complement IR4, -5 and -7 mutants respectively ((272), as reviewed in (234, 273)). Human chromosome 22q13 has been reported to restore radioresistance to IR6 mutants (250, 274). The genes complementing the DNA repair defects are referred to as X-ray repair cross complementing and receive the designation XRCC followed by the number of the complementation group for which they rescue (274).

The genes encoding IR5 mutant protein and Ku80 localize to the identical region on human chromosome 2q33-35 (275, 276). Indeed, DNA end binding activity characteristic of the Ku protein (see below) is absent from IR5 mutants (268, 277, 278). Not surprisingly, these IR5 mutants do not express the Ku heterodimer (267, 277, 279). Moreover, both DSB repair and V(D)J joining is restored to almost wild type levels upon transfection of Ku80 cDNA into these lines (268, 278-282). Taken together, these findings suggested that Ku80 is defective in IR5 mutants.

IR6 mutant, sxi-1, exhibits comparable IR sensitivity, V(D)J recombination and DNA end binding deficiencies to the IR5 mutants (see Table 1-1) (266). The XRCC6 gene has yet to be identified, however the similarities to IR5 mutants suggest that the XRCC6 protein is involved in the DNA-PK repair pathway (see below).
### TABLE 1-2

Summary of Mammalian DSB Repair Mutant Complementation Groups

<table>
<thead>
<tr>
<th>Complementation Group</th>
<th>DSB Repair Mutants</th>
<th>Chromosomal Locus</th>
<th>Complementing Gene/Product</th>
<th>V(D)J Recombination Activity</th>
<th>DNA End Binding Activity</th>
<th>DNA-PKcs Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Signal Joints</td>
<td>Coding Joints</td>
<td></td>
</tr>
<tr>
<td>IR4</td>
<td>XR-1</td>
<td>5q13-14</td>
<td>XRCC4</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>IR5</td>
<td>xrs-1 to xrs-6; XR-V15B; XR-V9B; sxi3; Ku80&quot; mice</td>
<td>2q33-35</td>
<td>XRCC5/Ku80</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IR6</td>
<td>sxi-1</td>
<td>?</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>IR7</td>
<td>V3; murine &amp; equine SCID; M059J</td>
<td>8q11</td>
<td>XRCC7/DNA-PKcs</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

* Equine SCID is the only member of IR7 to be defective in both signal and coding joint formation. See text for details.
The IR4 is the only DSB repair complementation group to possess both DNA end binding and DNA-PK kinase activities (8, 272). The IR4 mutant, XR-1, contains a unique cell cycle-dependent radiosensitivity during G1 to early S phase, but is radioresistant in late S phase (283). The XRCC4 gene maps to human chromosome 5q13-14 (284), and encodes a novel ubiquitously expressed protein that is highly hydrophilic (272). Interestingly, the XRCC4 protein contains a number of phosphorylation motifs, including a DNA-PK target site, as well as a putative nuclear localization signal (272). Transfection of XRCC4 cDNA restores both V(D)J recombination and DNA DSB repair to almost wild type levels (272).

- The DNA-dependent protein kinase

The DNA dependent protein kinase (DNA-PK) is a ubiquitously expressed serine/threonine kinase that is only activated upon binding to discontinuities in the DNA double helix (11, 13, 285-288). DNA-PK is comprised of two subunits, the DNA-binding Ku heterodimer, and a 460 kDa catalytic subunit, DNA-PKcs.

DNA-PK can phosphorylate a number of DNA-binding transcription factors in vitro, and that this activity was most efficient when DNA-PK co-localized to the same DNA molecule as its substrate (11, 13, 285, 286). These findings suggest that DNA-PK is involved in transcriptional control of DNA repair mechanisms and/or activation of DNA damage control checkpoints. DNA-PK catalytic activity is regulated (at least in vitro) by autophosphorylation of the DNA-PKcs which causes its dissociation from the Ku/DNA complex and renders the kinase inactive (289). Secondly, DNA-PK activity appears to be regulated in vivo by the specific proteolytic cleavage of DNA-PKcs by interleukin-1β converting enzyme (ICE) family members upon apoptotic stimuli (290-293).
The Ku complex is an abundant nuclear protein first identified as an autoantigen recognized by sera from patients with rheumatoid autoimmune disease (234), and references within. Ku is composed of two subunits of approximately 70 and 80 kDa, referred to as Ku70 and Ku80 respectively. Cells deficient in Ku80 protein lack detectable levels of Ku70, indicating that heterodimer formation is critical for the stability of the individual Ku proteins (277, 278, 294).

The Ku heterodimer is distinguished by its high affinity for double-stranded (ds) DNA ends, including hairpins (13, 278, 288, 295, 296), as well as single-stranded (ss) to ds DNA transitions (297). Indeed, Ku was later shown to be the component of the DNA-PK responsible for conferring the kinase's specificity for anomalies in the DNA helix (286, 287). Moreover, Ku80-deficient IR5 mutants lack DNA-PK activity and are defective in both V(D)J joining and DSB repair (268, 278-282). All these features are restored upon complementation with the XRCC5 gene/cDNA (268, 278-282), but not Ku70 cDNA (268, 278). This finding is consistent with reports suggesting that Ku represents the predominant, if not only, mode of activating DNA-PK (280).

Several observations suggest that the Ku proteins likely mediate other functions that are independent of DNA-PKcs. First, unlike the scid defect, Ku deficient cells are compromised in both signal and coding joint formation (146, 263, 266-268). Moreover, the signal joins isolated are predominantly imprecise, involving large deletions (146, 263), suggesting that Ku acts to protect DSBs from endonucleolytic cleavage. However, analysis of Ku80 deficient mice revealed that intact coding and signal ends accumulate in lymphoid cells, arguing that Ku does not function to protect broken DNA ends (298). Unexpectedly, Ku80 deficient mice are dwarfed, and their fibroblasts exhibit prolonged doubling
times and abnormal loss of proliferating cells (298, 299). This phenotype implicates the Ku proteins in growth control. Consistent with this notion, Ku is required to maintain normal telomere lengths in *Saccharomyces cerevisiae* (300).

- **The SCID Defect is DNA-PKcs**

Ku's involvement in both DNA repair and V(D)J recombination gave rise to speculation that DNA-PKcs might also be involved in these processes. Consequently, IR4, and IR7 mutants were screened for deficiencies in DNA-PK activity. Both IR4 and IR7 mutants contain normal Ku-mediated DNA end binding activity (267, 277), however only IR7 mutants lack DNA-PK activity (8, 9, 12, 301). Indeed, the level of DNA-PKcs expressed in both SCID and V3 IR7 mutant cells is reduced 10-102 fold (8-10, 301, 302). Thus, DNA-PKcs was identified as a strong candidate protein responsible for the *scid* defect.

Confirmation of this hypothesis was achieved by studies that mapped the SCID gene to the same region of human chromosome 8 that encodes the DNA-PKcs gene (303, 304). Introduction of human chromosome 8 fragments or yeast artificial chromosomes bearing the DNA-PKcs gene restores both DSB repair and V(D)J recombination processes to SCID cells (8-10, 302-304). Moreover, purified DNA-PKcs can restore DNA-PK kinase activity to extracts derived from SCID cells (8). Interestingly, equine SCID, which is only member of the IR7 group that is defective in both signal and coding joint formation (270), has no detectable DNA-PKcs protein (271).

**Characterization of the SCID Protein**

Clearer understanding of how DNA-PK functions came from the complete sequence analysis of the human DNA-PKcs gene. The carboxy-terminal 380 AA of the XRCC7 gene have strong homology with the phosphatidylinositol-3 (PI3) kinase superfamily (305, 306). DNA-PKcs belongs to PI3 kinase family subclass II, whose members are typically greater than 270 kDa, and share a common a
region of homology located in the extreme end of carboxy terminus that is lacking in class I PI3 kinase members. A member of this group includes the product of the gene mutated in AT (ATM). Generally PI3 kinase subclass II proteins are involved in controlling progression into cell cycle, a function which in turn affects DNA synthesis and cell cycle arrest in response to DNA damage. SCID cells do not exhibit any known perturbation in cell cycle control, especially in response to ionizing radiation (207, 259-262). No lipid phosphorylation activity has been observed for any member of PI3 kinase subclass II family, including DNA-PK (305). However, these kinases may require different substrates and/or conditions than described for the more conventional subclass I kinases for activity.

In wild type cells, DNA-PKcs protein is found in both cytosolic and nuclear extracts, however the tightly bound nuclear fraction of DNA-PKcs is not observed in murine SCID cells (301). Moreover, SCID DNA-PKcs protein levels are reduced 10-10² fold compared to wild type mice, and that this relative absence likely accounts for part of the SCID phenotype (8-10, 301). SCID cells contain normal amounts of steady-state DNA-PKcs mRNA, indicating that the decline in DNA-PKcs protein is likely a consequence of protein instability (301, 307). Insight into the genetic nature of the SCID defect has been provided by the identification of a single base pair mutation that results in a premature stop codon within the C-terminal kinase domain (301, 307). This mutation is predicted to truncate the protein by 83 residues (301, 307). Importantly, this deletion does not appear to affect the mutant DNA-PKcs's capacity to bind DNA, indicating that its association with the Ku antigen is not abrogated (301). Moreover, the nonsense mutation lies 100AA downstream of the conserved motifs required for kinase activity of conventional PI3 kinase family members (308), thus giving rise to speculation that a low level of catalytic activity may be
retained (301). However, no catalytic activity has been observed for the mutant DNA-PKcs in *in vitro* kinase assays, but this failure may reflect the sensitivity of the assays used (301). Collectively, these data can account for the 15-30% of leaky SCID mice (1).

Interestingly, almost wild type levels of the mutant DNA-PKcs has been observed in SCID/bcl-2 Il-7 expanded progenitor B cells (309). Moreover, the relatively abundant protein is readily found in nuclear extracts. This apparent discrepancy may be accounted for in part by the overexpression of *bcl-2* in these cells. The Bcl-2 protein has recently been shown to indirectly prevent the release of proapoptotic ICE family members from the mitochondria, subsequently preventing their activation (as reviewed in (310)), thereby accounting for the relatively high levels of DNA-PKcs in these cells. The inability to detect tightly bound nuclear associated DNA-PKcs in non-*bcl-2* transgenic SCID cells may in part reflect the lack of p460 present in these cells.

• **Role of DNA-PK in DSB Repair**

Insights into the role DNA-PK plays in DSB repair can be derived from the differential capacity of Ku- and DNA-PKcs-deficient cells to undergo V(D)J recombination. Ku-deficient cells exhibit a severe reduction in the frequency of both coding and signal joint resolution (146, 263, 278, 279, 298, 299), while coding joint formation is selectively affected in SCID mice (38). Moreover, the rare signal joins arising in Ku deficient cells exhibit large deletions (146). These findings gave rise to speculation that Ku functioned to stabilize and/or protect the DNA ends from endonucleases (8, 146, 267, 268, 277, 280). However, Ku80-deficient mice have normal levels of both signal and coding end intermediates, which do not exhibit any deletions from the original germline sequence (298). Moreover, full length signal ends accumulate at similar levels to wild type in Ku80-deficient cell lines transfected with RAG expression vectors and plasmid
recombination substrates (311). Thus, Ku does not appear to be involved in protection of DNA ends. Rather, Ku may act to disrupt or alter the cleavage synaptic complex, subsequently recruiting DNA-PKcs and other components of the DNA repair machinery necessary to repair the DNA lesions (298). Indeed, Ku has an ATP-dependent DNA helicase activity, which is thought to render DNA ends accessible for processing (312). The critical role DNA-PKcs plays in recruiting repair enzymes is underscored by the difference in severity of V(D)J joining between the murine and equine SCID mutations which can be explained by the complete lack of DNA-PKcs in equine SCID (270, 271).

Nonetheless, the exact role of DNA-PKcs in DSB repair has yet to be established. DNA-PKcs has the capacity to phosphorylate several transcription factors including SP1 (285), RNA polymerase I (313), as well as p53 (11) in vitro. Thus, putative roles for DNA-PK include signaling in the presence of DNA damage, recruitment of other components of the DNA repair pathway, and inhibiting transcription in the vicinity of the DNA discontinuities (234). DNA-PKcs is also thought to serve as a scaffold for assembling components of the DNA repair machinery in the vicinity of the DNA damage (305).

Interestingly, a DSB repair independent role for DNA-PK has recently been discovered (262). DNA-PK activity in wild type cells peaks at the G1/early S phase, and again at the G2 phase in wild type cells. Cells deficient in DNA-PK exhibit hypersensitivity to ionizing radiation only in the G1/early S phase, which correlates with a deficiency in DSB repair. However, both SCID and Ku80 deficient cells arrested in G2 demonstrated DSB repair activity indistinguishable from wild type cells. DNA-PK deficient cells subjected to irradiation progress to the G2 phase where they arrest, but fail to progress back into cell cycle. These results suggest that DNA-PK activity may directly or indirectly regulate the G2 checkpoint of DNA-damaged cells. Moreover, these results give rise to
speculation that DNA-PK may be involved in other cellular functions indirectly related to DSB repair and/or V(D)J recombination.
THESIS SUMMARY

The objective of my thesis is to understand how V(D)J rearrangements occur in the presence of the mutant SCID protein. My goal was to determine the extent to which Ig rearrangement occurs in SCID mice, and to assess the nature of these joins in order to gain knowledge into how these structures arose. The studies presented here were initiated prior to the identification of the SCID gene product, DNA-PKcs. The results of my thesis will be presented in the following chapters in context of the current understanding regarding DNA-PKcs and its role(s) in V(D)J recombination, DSB repair, and B cell development.

CHAPTER II:

This chapter provides a detailed account of the experimental procedures used in subsequent chapters.

CHAPTER III:

Chapter III provides a detailed assessment of endogenous DJH rearrangement in BM of non-leaky SCID mice. The frequency of resolved SCID DJH junctions is 1-10% of wild type BM. No aberrantly sized rearrangements were observed, despite the capacity of the PCR to amplify such junctions. Qualitative analysis reveals that these joins have fetal-like properties, such as restricted diversity, lack of N addition and joining by homology. Large P insertions characteristic of T lymphomas derived from leaky SCID mice are not observed in these joins.

CHAPTER IV:

A detailed assessment of VDJH and κ rearrangement in non-leaky SCID mice is presented. Non-aberrantly sized VDJH and κ rearrangements occur at 0.1-1% that of wild type B220⁺sIg⁻ cells. Aberrantly sized rearrangements are
observed only in the κ locus. The VDJH rearrangements exhibit normal number of deletions, and appear to be selected as 86% of the unique joins analyzed are productively rearranged (i.e. these joins are in frame, contain critical residues and lack stop codons). P insertions are found in these VDJH rearrangements. At least 40% of these structures, and most of the aberrant κ rearrangements involve joining by homology. Some of the κ joins exhibit N addition, suggesting that they arose during pro-B cell stage of development where TdT expression is still on.

CHAPTER V:

Chapter V provides a summary of the data presented in this thesis, and interprets these results in light of the recent advances in our understanding of the requirements for V(D)J recombination/DSB repair to occur.
CHAPTER II

METHODS

This chapter provides a summary of the various experimental techniques used throughout my thesis. Experimental procedures are grouped together by similarity of outcome, i.e. all DNA extraction protocols are located under the same subsection. Contents of this chapter have been adapted from "PCR assays for endogenous V(D)J rearrangement." by J. L. Pennycook, A. J. Marshall, & G. E. Wu, Immunology Methods Manual: The comprehensive sourcebook of techniques, I. Lefkovits (ed.). Academic Press. London. (1997).

I performed all the techniques listed in this chapter, with the exception of the cell sorting which was performed by D. Bouchard and S. Xie.
ISOLATION AND PURIFICATION OF BONE MARROW CELL DNA

MICE AND CELL LINES

The C.B-17 SCID (referred to in the thesis as SCID mice) and C.B-17 wild type (referred to in the thesis as C.B-17 mice) were originally obtained from Melvin Bosma (Fox Chase Cancer Center, Philadelphia, PA). Both mouse strains were bred and maintained in the animal facility of the Ontario Cancer Institute under defined flora conditions (Toronto, ON, Canada). SCID mice were screened for serum Ig by ELISA (see below); none of the SCID mice used had detectable serum Ig. BM and spleens were isolated from individual mice at 6 to 8 weeks of age.

A-MuLV lines used for PCR standards were generated by M. Atkinson. Monoclonal Ab producing cell lines 14.8 and 3360 were gifts of C. J. Paige. MC-9 is a mast cell line generously provided by S. A. Berger.

DETECTION OF SERUM IMMUNOGLOBULIN

The enzyme-linked immunoassay (ELISA) was used to screen for non-leaky SCID mice. ELISA plates (Costar) were coated with 50 μl/well of 5 μg/ml of affinity pure goat anti-mouse IgM, μ chain specific (Jackson ImmunoResearch Laboratories, Inc.) diluted in 0.05M Tris pH 9.5, 0.15M NaCl buffer. The plates were incubated for 1 hour at 37°C, and were washed 3X with tap water. Wells were blocked with 50 μl/well PBS with 5% fetal calf serum (FCS) for 1 hour at 37°C. The ELISA plates were washed 6X with tap water.

Serum was obtained by tail bleeding SCID mice without the use of heparin to permit the blood to coagulate such that the sera separated. The coagulated blood samples were centrifuged gently at 400g to enhance
further separation of the sera without disrupting the pelleted red blood cells (RBC). A ten-fold serial dilution of the SCID sera was made in PBS with 5% FCS, and 50μl of each dilution was then added to the plate in duplicate. Each plate contained a negative control (i.e. PBS with 5% FCS alone) and a serial dilution of purified mouse IgM (Jackson ImmunoResearch Laboratories) which was used as a quantitative standard to determine the amount of serum IgM present in the sera. The purified IgM standards were also diluted in PBS with 5% FCS, and of 50μl each dilution was used. Plates coated with sera samples were incubated for 1 hour at 37°C, and were washed 6X with tap water.

Affinity purified goat anti-mouse IgM (μ-chain specific) peroxide conjugate (Sigma ImmunoChemicals) was diluted 1:1000 in PBS + 5% FCS, and 50μl was added to each well. The plates were incubated and washed as before. Detection of serum Ig was made possible by adding 0.5 mg/ml 2,2'-azino-bis (3 ethylbenzthiazoline) 6-sulfonic acid (ABTS) (Sigma) diluted in 0.05M phosphate-citrate buffer with 0.03% sodium perborate (Sigma). ABTS is a substrate for the peroxide conjugated to the goat anti-mouse IgM Ab and its breakdown products result in a colour change. The plates were incubated for 10 minutes at 37 °C to allow for the ABTS substrate to be catalyzed. Plates were read at 405/630 O.D. in a Molecular Devices Thermo max spectometer.

PURIFICATION OF CELLS

BM was aspirated from femurs and tibias of individual mice using PBS + 5% FCS in a 26G needle syringe. The aspirated cells were centrifuged at 350g for five minutes at 4°C. In order to lyse the erythrocytes, the resulting pellet was resuspended in an ACK lysis
buffer (0.155M ammonium chloride, 0.1mM disodium EDTA, 0.01M potassium bicarbonate) at 10^8 cells/ml and was incubated on ice for three minutes. Cells were then washed three times in a 12 ml volume of PBS + 5% FCS at 350g for five minutes at 4°C. Whole BM and B220^+ cell suspensions were prepared for individual SCID mice, while C.B-17 cells were pooled. BM was purified by either panning and/or cytofluorometry.

Cell preparations used in cytofluorometry were stained with RA3-6B2 anti-B220-PE (rat IgG_2a, κ) and R6-60.2 anti-IgM-FITC (rat IgG_2a), or the appropriate isotypic control antibodies (PharMinigen, San Diego, CA). Cell preparations consisted of either BM cells directly isolated or enriched for B220^+ cells. All cell preparations were stained at 10^8 cells/ml diluted in PBS, 5% FCS. Preparations stained for IgM expression had 0.5 µg/ml of anti-IgM-FITC added. Samples stained for B220 expression had 0.2 µg/ml of anti-B220-PE Ab. Normal whole BM was always used to set the parameters necessary for sorting cells. These controls involved staining cells with the above mAbs either alone or together. Cells were also stained with the appropriate isotype-matched control Abs in order to control for non-specific staining. SCID cell preparations were always double stained for B220 and IgM expression. Isotype controls were used in identical concentrations used for the specific Abs for which they controlled background staining. Staining occurred for 30 minutes at 4°C in the dark. The cells were then washed three times in three ml of PBS, 5% FCS, and centrifuged at 350g for five minutes at 4°C each time. Stained cells were separated using FACStar flow cytometer (Becton Dickinson & Co., Mountain View CA).
DNA PREPARATION

Three methods of DNA purification were utilized in the duration of my thesis. The traditional phenol/chloroform method and the direct cell lysis methods of DNA extraction were used to isolate genomic DNA. Generally, the phenol/chloroform method was used when > $10^7$ cells were available for analysis, while the direct cell lysis was used when cell numbers were limiting. The miniprep protocol was used to quickly isolate plasmid DNA for sequence analysis.

- **Phenol/Chloroform Extraction**

  Cells were gently harvested at 350g in 4°C centrifuge for one minute, and were washed twice with PBS under the same centrifugation conditions in order to remove all traces of FCS. Cells were then resuspended in TNE buffer (10 mM Tris pH 8.0 Ultra pure, ICN; 150 mM NaCl Sigma; 1 mM EDTA Analar, BDH Inc.) at a concentration of 4x10⁷ cells/ml (314). An equal volume of TNE + 0.4% SDS buffer was then added to the cell suspension, resulting in a final concentration of 2x10⁷ cells/ml. Proteinase K was added to a final concentration of 100 mg/ml. The cell lysates were incubated for two hours at 56°C to prevent DNase activity during lysis while allowing for efficient protein degradation.

  DNA was extracted with an equal volume of phenol equilibrated with Tris pH 8.0, 1mM EDTA: samples were allowed to rotate at 4°C in the dark for ten minutes and were then centrifuged for ten minutes at 15 000g at room temperature (RT) to recover the aqueous phase. The aqueous phase was carefully removed and transferred to a fresh tube. It was then extracted with an equal volume of phenol chloroform isoamyl alcohol (1:1) (PCI) using the same conditions. A final chloroform isoamyl alcohol (24:1) extraction was performed when there was no visible protein interface after centrifugation. The final aqueous phase was transferred to a fresh tube and 0.1 volume of 3M NaAc pH
5.2, and two volumes of ice cold 100% EtOH were then added. The tubes were mixed gently by inversion. DNA was either spooled using a sealed pasture pipette (315), or was retrieved as a pellet using the identical centrifugation conditions described earlier. The isolated DNA was immediately washed in cold 70% EtOH, and allowed to air dry for one minute. The genomic DNA was then dissolved in TE buffer, and stored at 4°C until required.

- **Direct Cell Lysis**

The second method I used to isolate genomic DNA was the direct PCR lysis protocol modified from (105). Isolated cells were washed three times in cold PBS at 2 600g at 4°C for two minutes per wash. The cells were resuspended at concentrations of either 10^3-10^4 cells/μl in PCR lysis buffer (10 mM Tris pH 8.3, 2 mM MgCl_2, 50 mM KCl, 0.45% Nonidet P40, 0.45% Tween 20) with 60 μg/ml Proteinase K. Cells were lysed for one hour at 56°C, followed by a 15 minute incubation 90°C in order to inactivate the proteinase K. Lysates were stored in small volumes at -20°C until needed.

- **Miniprep DNA Extraction**

The miniprep DNA extraction protocol was utilized to isolate small quantities of plasmid DNA (<10μg) for sequencing analysis. A modified alkaline miniprep protocol was utilized. 1.5 ml of inoculated overnight bacterial cultures were harvested by centrifuging briefly at 6 600 g. Pelleted bacterial cells are resuspended in 200 μl of isotonic buffer (50mM glucose, 25mM Tris pH 8.0, 10mM EDTA), and lysed by adding 400 μl of alkaline lysis buffer (0.2N NaOH, 1% SDS). Bacterial protein and chromosomal DNA were precipitated by adding 300 μl of 7.5M NH_4Ac, gently mixing by inversion, and incubating for ten minutes on ice. The DNA preparations were centrifuged at 15 000 g for ten minutes at RT. 800 μl of supernatant were removed and transferred to a new tube, where 0.8 volumes of isopropanol were added to precipitate the plasmid
DNA. The mixture were then centrifuged at 15 000 g for ten minutes at 4 °C. The supernatant was aspirated, and the pelleted DNA was washed in 70% ice cold ethanol. The resulting pellets were dissolved in 50 μl of TE buffer with 100 μg/ml RNase A to digest any contaminating RNA.

Diagnostic restriction digestion with EcoR1 was performed on the resulting minipreps to confirm the presence of a PCR product, and to estimate the quantity of DNA present. Most diagnostic digests were Southern blotted and probed with Ig loci specific probes.

**PCR ASSAYS**

The DNA based PCR methods take advantage of the changing structure of the Ig loci during the recombination process. When the Ig loci are recombined during V(D)J recombination, genetic elements which are separated by large distances in the germline locus are brought together to create a new DNA structure. Thus, specific PCR primers for these genetic elements will only amplify the recombined structures. Amplification strategies are presented in Figures 2-1 to 2-3. Summaries of PCR primer sequences and conditions are provided in Tables 2-1 and 2-2 respectively.

**D-JH REARRANGEMENT**

The PCR strategy for D-JH joining is depicted in Figure 2-1, and the sequences for all the primers used are detailed in Table 2-1. The DSF primer is a single oligomer that hybridizes to the 5' RSS of all Dsp and DFL elements (26). There are three mismatches between the DSF primer and the 5' RSS of the DH elements, with the exception of Dsp 2.5, which has four. For all Dsp and DFL elements, two mismatches are located in tandem at positions seven and eight of
**Figure 2-1:** Schematic diagram of the DJH PCR strategy used in Chapter III, showing the oligomer positions and the expected PCR products. Refer to text and to Tables 2-1 and 2-2 for primer sequences and details of PCR conditions. This diagram is not drawn to scale.
Figure 2-2: Schematic diagram illustrating the VDJH PCR amplification strategy used in Chapter IV. Oligomer primer sites and PCR product sizes are indicated. Details of PCR primers and conditions are provided in the text and in Tables 2-1 and 2-2.
Figure 2-3: Schematic diagram of the \( \kappa \) PCR strategy diagram employed in Chapter IV. Oligomer binding sites and sizes of expected PCR products are indicated. Refer to text and Tables 2-1 and 2-2 for details regarding primer sequences and PCR conditions.
the primer sequence, disrupting the AT-rich tract of the nonamer sequence, and contributing to an embedded BamHI site that facilitates cloning. The third mutation site for all Dsp elements lies at position 13 of the primer, whereas positions five and one are altered for DSF 16.1 and DSF 16.2 elements respectively. The fourth mismatch site in Dsp 2.5 lies at position 18 of the primer.

The JH4 primer is a 26mer that hybridizes to the DNA sequence immediately adjacent to the JH4 coding element in the JH-C intron (26), see Table 2-1. It has one internal mismatch at position nine to facilitate cloning. The JH4-IN 26mer hybridizes to sequences contained entirely within JH4; it shares no overlapping homology with JH4 (245). The JH2 26mer hybridizes ~210bp downstream of JH2 in the JH2-JH3 intergenic sequence (245). Neither DSF/JH4 nor DSF/JH2 primer pairs amplify VDJH rearrangements (26, 245). Both the JH4-IN and JH2 primers were utilized in semi-nested secondary PCR amplifications. The JH4-IN2 26mer hybridizes within JH4 but shares no homology with JH4-IN, it is used mainly as a probe (52). The JH2-IN 26mer hybridizes ~170 bases 3' of JH2 in the JH2-JH3 intergenic sequence and is used primarily as a probe.

PCR conditions used to amplify DJH rearrangements are detailed in Table 2-2 (also see PCR Assay Conditions). Simultaneous amplification of DNA derived from cell lines bearing Dsp and DSF DJH rearrangements in a 1:1 ratio revealed no inherent DH family amplification biases (26). Titration standards derived from four A-MuLV lines of known DJH rearrangement status (i.e. involving JH1 to JH4 elements) were used to permit the determination of the number of DJH structures present in 1 μg of DNA derived from SCID or C.B-17 whole BM. Briefly, equal numbers of DJH targets for each JH element were used because amplification of different sized PCR DNA fragments occurs at varying efficiencies (26). DNA from the four standard A-MuLV lines was mixed in equal
amounts, and serially diluted into DNA derived from an VDJH/VDJH rearranged A-MuLV line (CB32.12), such that a constant amount of DNA was amplified in each member of the titration standard. The titration standard was amplified in parallel to the experimental samples.

Semi-nested secondary PCR amplifications involved a 10-10² fold dilution of the sample of interest before being subjected to the identical PCR cycling conditions used for primary amplifications (245).

V-DJH REARRANGEMENT

The VHall 22mer hybridizes to the sequence corresponding to AA 1-9 of 80% of all VH elements (316). The VHall primer contains five sites of degeneracy (see Table 2-1), and is thus a mixture of different oligonucleotides. The identical JH4 and JH2 primers described above were used to detect VDJH rearrangements (see Table 2-1 for sequences and Table 2-2 for PCR conditions). The complexity of the VH loci made it difficult to determine whether certain VH elements and/or families are preferentially amplified by the VHall primer, thus no adequate titration standard similar to the one used in Chapter III could be created.

Semi-nested secondary PCR amplification using the VHall primer in conjunction with either JH4-IN or JH2 primers results in a high amount of nonspecific PCR products. In order to enhance the fidelity of the amplification, three VH family specific primers internal and/or overlapping with the VHall primer were utilized (see Table 2-1). The VHJ558 34mer hybridizes in third framework region of VHJ558 family gene segments (105). The VHQ52 34mer hybridizes in third framework region of VHQ52 family gene segments (105). The VHJ558 and VHQ52 primers were mixed in equimolar amounts and used to amplify VHall/JH4 amplified DNA samples. The VH7183 32mer hybridizes to sequences corresponding to AA five to eleven of the VH7183 family gene segments
<table>
<thead>
<tr>
<th>Oligo</th>
<th>Sequence (5'-&gt;3')</th>
<th>Oligo Size</th>
<th>Locus</th>
<th>Region of Hybridization</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSF (5')</td>
<td>AGG GAT CCT TGT GAA GGG ATC TAC TAC TGT G</td>
<td>31mer</td>
<td>D-&gt;JH</td>
<td>5' RSS</td>
</tr>
<tr>
<td>JH4 (3')</td>
<td>AAA GAC CTG CAG AGG CCA TTT TTA CC</td>
<td>26mer</td>
<td>D-&gt;JH</td>
<td>J-C intron just 3'JH4</td>
</tr>
<tr>
<td>JH2 (3')</td>
<td>TGG CCA GGA TCC CTA TAA ATC TCT GG</td>
<td>26mer</td>
<td>D-&gt;JH</td>
<td>210bp 3' of JH2 in JH2/JH3 intervening sequence</td>
</tr>
<tr>
<td>JH2-IN (3')</td>
<td>CTG AGG ATG TCT TGC ATC AGC CA</td>
<td>26mer</td>
<td>D-&gt;JH</td>
<td>170bp 3' of JH2 in JH2/JH3 intervening sequence</td>
</tr>
<tr>
<td>JH4-IN (3')</td>
<td>GAG GAG ATG TCT TGC ATC AGC CA</td>
<td>26mer</td>
<td>D-&gt;JH</td>
<td>JH4 element, does not overlap with JH4 3'</td>
</tr>
<tr>
<td>JH4-IN2 (P)</td>
<td>ACC CCA GTA TGC CAT AGC GTA AT</td>
<td>26mer</td>
<td>D-&gt;JH</td>
<td>JH4 element, does not overlap with JH4-IN 12-36 of VH segment coding sequence</td>
</tr>
<tr>
<td>VH7183 (5')</td>
<td>CGC GCG GCC GCG TGG AGT CTG GGG GAG GCT TA</td>
<td>32mer</td>
<td>V-&gt;DjH</td>
<td>Framework 3</td>
</tr>
<tr>
<td>VH3558 (5')</td>
<td>CGA GCT TCC CA(A/G) CAC AGC CT(A/T) CAT GCA (A/G) CT CA(A/G) C</td>
<td>34mer</td>
<td>V-&gt;DjH</td>
<td>Framework 3</td>
</tr>
<tr>
<td>VHQS52 (5')</td>
<td>CGG TAC CAG ACT GA(A/G) CAT CA(C/G) CAA GGA CAA (C/T) TCC</td>
<td>34mer</td>
<td>V-&gt;DjH</td>
<td>Framework 3</td>
</tr>
<tr>
<td>Vxcon (5')</td>
<td>GGC TGC AGC (C/G) TTC AGT GGC AGT GG(A/G) T(C/A/T) G(G/A) AC</td>
<td>32mer</td>
<td>Kappa</td>
<td>Framework 3</td>
</tr>
<tr>
<td>Vxcon-S (5')</td>
<td>GGC TGC AA(C/G) TTC AGT GGC AGT GG</td>
<td>23mer</td>
<td>Kappa</td>
<td>Framework 3</td>
</tr>
<tr>
<td>Jx2 (3')</td>
<td>GGT TAG ACT TAG TGA ACA AGA GTC GGG AG</td>
<td>29mer</td>
<td>Kappa</td>
<td>3' of Jx2</td>
</tr>
<tr>
<td>Jx2-IN (3')</td>
<td>CCA AGC TTT CCA GCT TGG TCC CCC CTC CGA A</td>
<td>31mer</td>
<td>Kappa</td>
<td>3' of Jx2 3' primer</td>
</tr>
<tr>
<td>Jx5 (3')</td>
<td>TGC CAC GTC AAC TGA TAA TGA GCC CTC TC</td>
<td>29mer</td>
<td>Kappa</td>
<td>3' of Jx5</td>
</tr>
<tr>
<td>Jx5-INM (3')</td>
<td>CTC CTC gCc TGA AAA CTT GTC TGT TAC</td>
<td>27mer</td>
<td>Kappa</td>
<td>5' of Jx5</td>
</tr>
<tr>
<td>Jx5-IN (3')</td>
<td>CTC CTA ACA TGA AAA CTT GTG TCT TAC ACA</td>
<td>30mer</td>
<td>Kappa</td>
<td>5' of Jx5 3' primer</td>
</tr>
<tr>
<td>Jx5-IN2 (P)</td>
<td>GTG TAC TTA CGT TCC ATG TCC AG</td>
<td>23mer</td>
<td>Kappa</td>
<td>3' Jx5 coding sequence</td>
</tr>
</tbody>
</table>

Brackets denote sites of degeneracy within the primer sequence. Sequences with asterisks or underlining within Vxcon denote regions of identity shared with Jx4 and Jx5 elements respectively. Lower case lettering within the Jx5-INM 3' primer indicate mismatches from the germline/Jx5-INM 3' sequence which disrupt sites of self-annealing and promote efficient PCR amplification. (P) denotes primers used as oligomer probes to detect amplified products. See text for details.
(317). Amplification with either the VHJ558 or VHQ52 primers generates PCR products ~200 bp smaller than VHALL- or VH7183/JH4 or JH2 primer pairs.

**κ REARRANGEMENT**

For analysis of the VJK rearrangements, either Vκcon or Vκcon-S was used with the Jκ5 primer (see Table 2-1). PCR was carried out under the same cycling conditions as cited for the VDJH assays (see Table 2-2). The Vκcon 32mer hybridizes in framework 3 of Vκ elements (104). Vκcon-S 23mer is identical to Vκcon from position one to 23, the final 9 nt were not synthesized in order to remove any possible sequence homology with the Jκ elements. The Jκ5 29mer hybridizes ~150 bases 3' of the Jκ5 element (318). The Jκ5-IN 30mer hybridizes approximately 20 bases 3' of the Jκ5 element in the Jκ-C intervening sequence (318). It was used primarily as a probe because it does not work well as a primer for PCR amplification. The Jκ5-INM is a 27mer. It hybridizes to the same location as Jκ5-IN, but contains three mismatches to allow for efficient annealing and subsequent PCR amplification. The Jκ2 26mer hybridizes 10 bases downstream of Jκ2 element (318), and is used primarily in secondary amplification. The Jκ2-IN primer hybridizes to the Jκ2 element itself (104), and is used primarily as a probe.

The large diversity in the number of Vκ elements made it difficult to create determine whether preferential amplification of an individual Vκ element and/or family occurs with the Vκcon primer. Consequently, no titration standard was created for this assay.

**PCR ASSAY CONDITIONS**

PCR was carried out in 10mM Tris pH 8.3, 50mM KCl, 1.8-2.1 mM MgCl₂, 0.5% Triton X-100, 100μg/ml BSA (Boehringer Mannheim), 200μM of each dNTP (Boehringer Mannheim), 0.5 μM of each primer (Gibco BRL), and 2.5 U of Taq
Table 2-2:
PCR Conditions for Detection of Ig Gene Rearrangements

<table>
<thead>
<tr>
<th>ASSAY</th>
<th>PRIMERS</th>
<th>FINAL [MgCl₂]</th>
<th>DENATURATION</th>
<th>ANNEALING</th>
<th>EXTENSION</th>
<th>Increase in extension (sec/cycle)</th>
<th>No. Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-J₇₇</td>
<td>DSF</td>
<td>JH4</td>
<td>2.5 mM</td>
<td>94°C 0:30</td>
<td>62°C 0:31</td>
<td>72°C 1:00</td>
<td>3</td>
</tr>
<tr>
<td>V-DJ₇₇</td>
<td>V₇HALL</td>
<td>JH4</td>
<td>1.8 mM</td>
<td>94°C 1:00</td>
<td>60°C 1:00</td>
<td>72°C 2:00</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>V₇H7183</td>
<td>JH4</td>
<td>1.8 mM</td>
<td>94°C 1:00</td>
<td>60°C 1:00</td>
<td>72°C 2:00</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>V₇H558</td>
<td>JH4</td>
<td>1.8 mM</td>
<td>94°C 1:00</td>
<td>60°C 1:00</td>
<td>72°C 2:00</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>V₇HQ52</td>
<td>JH4</td>
<td>1.8 mM</td>
<td>94°C 1:00</td>
<td>60°C 1:00</td>
<td>72°C 2:00</td>
<td>5</td>
</tr>
<tr>
<td>V-Jk</td>
<td>V₇kcon/</td>
<td>Jk5</td>
<td>1.8 mM</td>
<td>94°C 1:00</td>
<td>60°C 1:00</td>
<td>72°C 2:00</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>V₇kcon-s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
polymerase (Boehringer Mannheim). For each assay, samples were prepared for PCR without Taq polymerase, and were incubated for a total of ten minutes at 85°C. During this incubation, 10μl of 2.5U of Taq polymerase in 1X Taq buffer was added to each sample. This incubation step is referred to as a "hot start" and is performed to minimize non-specific annealing and extension of the PCR primers (319). The hot start was followed by 30 cycles of the appropriate cycling condition for the given assay (see Table 2-2). The extension rate of the Taq polymerase under the conditions outlined in Table 2-2 is approximately one kb per minute. The limitations in the extension rate is reflected in the lower intensity of the larger PCR products (see Figures 3-1, 4-1, 4-4). Each successive PCR cycle involved an increase in the extension phase of PCR by either three or five seconds (see Table 2-2). This increase is referred to as autocycling and is performed in order to compensate for the ever decreasing efficiency of the Taq polymerase as it proceeds through successive rounds of heating and cooling. A ten minute extension period at 72°C followed the 30 cycles. This extension period allows the completion of any partial amplifications, and ensures that the Taq polymerase dependent non-templated addition of a single deoxyadenosine to the 3'-ends of duplex PCR products occurs (see Cloning and Sequencing below).

In order to facilitate cloning, primary amplifications from SCID samples were diluted 10- and 10²-fold prior to secondary PCR amplification. Secondary semi-nested DJH amplifications were carried out with the DSF primer and either JH4-IN or JH2 primers, under identical cycling conditions as the primary amplification (Table 2-2). The VHALL primer is too degenerate for efficient secondary amplification, thus secondary amplification was carried out using either the VHJ558 and VHQ52 primers together, or VH7183 primer with the JH2 or JH4-IN nested primers under the same cycling conditions as VHALL/JH4. For cloning of SCID VJκ rearrangements, secondary amplifications were carried out on 10- and 10²-fold
dilutions of primary PCR product with Vkcon or Vkcon-S primers and either Jk5, Jk2 or Jk5-INM primers. Cycling conditions were identical to those described for the VDJH analysis (see Table 2-2).

Alpha actin primers (α-Actin 5' 5'-GACATGAGAAGATCTGGCACCACACAC-3' and α-Actin 3' 5'-CGCACAATCTCAGGCTCCAG-3' (76)) were used in order to control for the number of cell equivalents used from sample to sample. Ten µl of VDJH or VJk amplified samples were directly subjected to a secondary PCR amplification with α-actin primers under the identical PCR cycling conditions described above.

DETECTION OF Ig REARRANGEMENTS

SOUTHERN BLOT ANALYSIS

15% of each PCR reaction was loaded on a 1.5% agarose gel (Sigma Chemical Co., St. Louis, MO) and electrophoresed in TAE buffer. Either Zeta probe (Bio-Rad Laboratories) or N-Hybond nylon membrane (Amersham, Chalfont, UK) were used for the Southern procedure. Briefly, agarose gels were denatured for 20 minutes in 0.4 N NaOH, and were neutralized in 1M Tris pH 7.5 for 20 minutes. The neutralized gel was then blotted using 10X SSC transfer buffer as described for Southern blotting (314). The transfer occurred overnight, followed by UV crosslinking (Fischer Scientific FB UVXL-1000) prior to hybridization.

PROBING

Southern blots were probed using either random primed probes, or labeled oligomer probes. Random primed probes were either radioactively labeled, or digoxigenin (DIG) labeled followed by detection using chemillumenescence.
Random Primed Detection

Random primed labeling exploits the Klenow enzyme's ability to polymerize DNA. Briefly, random hexanucleotides hybridize to denatured DNA fragments, and the 3' -OH termini serve as primers for DNA synthesis using Klenow. Modified deoxynucleoside triphosphates (αP32–dCTP or DIG-11-dUTP) are incorporated into the newly synthesized complementary strand.

Radioactive Detection

Filters to be analyzed are prehybridized in 50% formamide, 1X Denhardt's solution (314), 5X SSC, 1% SDS, 100 µg/ml denatured herring sperm DNA. 100 ng of DNA fragment in 37 µl H2O was boiled for 10 minutes, and placed on ice. The denatured DNA was labeled by adding 10 µl of 5X OLB buffer, 5 µl αP32–dCTP and Klenow (10 units/µl) for 2-3 hours 37°C. The OLB is made by adding 10 µl Solution II [1 ml of Solution I (1.25M Tris pH 8.0, 0.125M MgCl), 18 µl mercaptoethanol, 5 µl each of 100mM dATP, dGTP, dTTP (Boehringer Mannheim)], 25µl Solution III [2M Hepes, pH 6.6], 15µl Solution IV [(dN)6 90 units/ml (Pharmacia)] and 5µl dH2O.

The labeled probe was purified on a G50 Sephadex spin column (Costar) which was made by twice adding 500µl of TE soaked G50 Sephadex and centrifuging at 700g for 30 seconds. The flow through was removed, and the labeling reaction added, and centrifuged as described. The purified probe was recovered, and it was boiled to denature it, and then directly added to the prehybridized blots.

Hybridization was carried out overnight at differing temperatures depending on the probe used. JH4 fragment probes (~900bp) were hybridized at a lower temperature (37°C) in order to detect rearrangements to the JH4 element because there is less than 50 bp of homology with the fragment. All other random primed probes were hybridized at 42°C.
After hybridization, filters probed with the JH4 fragment probe were washed twice in 2X SSC, 0.1% SDS at RT for five minutes with shaking. Two more stringent washes were carried out at 42°C for 30 minutes in 1X SSC, 0.1% SDS. Filters probed with other random primed probes were washed at 65°C in 0.1X SSC, 0.1% SDS for 30 minutes.

Non-Radioactive Detection

Non-radioactively labeled probes were used primarily to detect colonies and plasmid DNA samples that harbored PCR inserts. This procedure was ideal in that a single labeled probe could be reused multiple times without the loss of sensitivity. DNA fragments are labeled using the same principle outlined above for radioactive labels probes, except that DIG-dUTP is used in lieu of αP³²-dCTP.

Blots to be screened were prehybridized for a minimum of one hour in prehybridization solution (5x SSC, 5% Skim milk powder (Carnation), 50% Formamide, 0.1% N Lauryl Sarcosine). As with radioactively labeled randomly primed oligos, hybridization temperatures varied depending upon the probe and product size desired to be detected (see above).

In a 30μl volume, 100 ng of denatured fragment DNA is labeled with DIG-dUTP according to manufactures' instruction (Boehringer Mannheim) using Klenow polymerase for one hour at 37°C. The reaction is stopped by adding 3 μl of 0.2M EDTA. The labeled DNA is then precipitated by adding 0.1 volumes of 4M LiCl and three volumes of 100% ice cold ethanol on dry ice for 30 minutes. The DNA was centrifuged for 10 minutes at 20 200 g; the resulting pellet was washed in 70% ice cold ethanol and centrifuged as before. The pellet was dried using a vacuum centrifuge, and was then resuspended in 50 μl TE and allowed to dissolve for 30 minutes at 37°C. The labeled oligo was added to 10 ml of prehybridization solution with the filters to be screened (see above). The probe
was boiled for 10 minutes prior to use, and was added to the hybridization container(s). Nitrocellulose membranes (Schleicher & Schuell) were hybridized at 42°C or 37°C overnight with DIG-dUTP labeled (Biochemica; Boehringer Mannheim) JH2 or JH4 probes respectively. Membranes were washed as described above for radioactively labeled random probes.

To detect the signal, filters were first blocked in a 10-20ml solution of 100mM Tris HCl, pH7.5, 150mM NaCl, 0.02% Tween 20 and 2% Skim milk powder (Carnation) for three hours. Anti-DIG horseradish peroxidase (POD) conjugated antibody (BMH) was diluted 10⁻³ into the blocking solution and incubated for 1 hour at room temperature with shaking. The blots were then washed four times in 50ml volumes of 100mM Tris HCl, pH7.5, 150mM NaCl, 0.02% Tween 20 for a total of 1 hour, and were then subjected to the Enhanced Chemiluminescence (ECL) detection system (Amersham) for 60 seconds. The blots were then exposed to film for 30-120 seconds. Positive colonies were identified and corresponding bacterial colonies were inoculated into 2XYT media for plasmid DNA extraction (see Miniprep protocol).

Radioactive Oligo Labeling:

Radioactively labeled oligos were used for quantitative analysis as well as for screening colony lifts. Filters were first prehybridized at 42°C in 5X SSPE, 5x Denhardt’s solution, 0.5% SDS for 1-2 hours and then the appropriate γ³²P-labeled oligonucleotides (see Table 2-1) was added for hybridization. 100ng of oligo was labeled using T4 polynucleotide polymerase (NEBlabs, Mississauga, ON) according to the manufacture’s instructions. Labeled oligos were added to prehybridized blots for 3 hours at 42°C, and then washed 2x in 2X SSC, 0.1% SDS for 5 minutes at 42°C, and then 2x in 2X SSC, 0.1% SDS for 30 minutes at 50°C. All blots were either exposed to NEN film with intensifying screens, and in a PhosphorImager cassette (Molecular Dynamics). Imaging and quantification was
done using a Molecular Dynamics PhosphorImager and ImageQuant 4.1 software.

**QUANTIFICATION**

Quantification was accomplished by drawing a grid around the bands of interest. Background values were defined by drawing a large rectangle across the expanse of the blot image being analyzed. Using the "Object Average" function of ImageQuant 4.1 program, only pixel readings above the background object's average pixel intensity value were included in the output reading for a particular grid section of interest.

**STATISTICAL ANALYSIS**

In Chapter III, parallel amplification of the titration standards with SCID and C.B-17 DNA allowed for the estimation of the number of DJH rearrangements present in 1μg of DNA derived from these two strains of mice. Consequently, an estimation of the total number of DJH structures in whole BM and spleen based on the curves generated by the titration standard values could be made for either strain using the TREND function of Excel 3.0 (Microsoft). The arithmetic mean, which is the sum of all measurements (i.e. the number of DJH targets) divided by the number of measurements made in a sample (i.e. number of mice) (see Equation #1 below) (320), was calculated for both strains of mice using StatviewII.

**Equation 1:** Arithmetic Mean

\[ \bar{X} = \frac{\sum X_i}{n} \]

The variation in the mean number of DJH rearrangements for each strain of mouse was calculated using the standard error (SE) equation. The SE measures the
precision with which a mean has been determined (see Equation #2 below) (320). The SE were presented in Chapter III were determined by using StatviewII.

**Equation 2:** Standard Error

\[
S_X = \sqrt{\frac{S^2}{n}} \quad \text{where} \quad S^2 = \frac{\sum (X_i - \bar{X})^2}{n - 1}
\]

No adequate titration standards could be derived to permit accurate quantification of the number of VDJH and κ rearrangements in both C.B-17 and SCID BM cells because of the complexity of the VH and Vk loci. Consequently, separate VDJH and κ amplifications could not be compared to one another.

Generally, DNA lysates from individual SCID and pooled C.B-17 BM cell populations were amplified in either triplicate or duplicate (see Figures 4-1, 4-4). The geometric mean was used to determine the relative frequency of VDJH and κ rearrangement in BM lysates derived from both SCID and C.B-17 mice. The geometric mean is a measure of central tendency and is calculated as the antilogarithm of the arithmetic mean of the logarithms of the data (see Equation 3) using Microsoft Excel 5.0 (320, 321). The geometric mean is used in instances when a series of measurements exhibit "a degree of positive skewness", which is removed by taking logarithms (321). The large variation among repeated samplings of the same population generally reflects the manner in which the measurements are made. Here, differences in the extent of VDJH or κ rearrangement in the initial cell lysate samplings are low, however such small variations can appear exponentially different after PCR amplification. Indeed, previous analysis determining the limitations of D-JH joining revealed that 10 DJH targets/μg of DNA was the lower limit of sensitivity of these assays. However, repetitive amplifications of the same titration standard of 10 targets for each of the JH elements resulted in variable detection of rearrangements to the four JH elements, i.e. not all of the JH element rearrangements were always amplified (26).
The standard deviation of relative intensity for rearrangements to each JH and Jκ element was determined for triplicate amplifications for each SCID mouse and the pooled C.B-17 BM cells within an individual experiment. The standard deviation is the positive square root of the variance as defined in the Equation 4 (320). The standard deviation for rearrangements in individual mice was calculated using Microsoft Excel 5.0.

**Equation 4:** Standard Deviation

\[
S = \sqrt{\left( \sum X_i^2 - \left( \frac{\sum X_i}{n} \right)^2 \right) / n - 1}
\]

**CLONING AND SEQUENCING**

One of two methods was used to clone the PCR amplified Ig products: TA cloning or the use of embedded restriction endonuclease sites in the PCR primers (VDJH and DJH products only). During the PCR reaction, a single deoxyadenosine (A) is added to the 3' ends of all duplex PCR products by the Taq polymerase. The TA cloning vector is a linearized plasmid that contains single 3' deoxythymine (T) overhangs, thus permitting the direct ligation of the PCR product without the need to purify and/or modify the ends of the product. PCR products from secondary amplifications were either directly ligated or diluted 10⁻¹ and a 1 μl aliquot was ligated into the TA Cloning vector (Invitrogen, San Diego, CA). Ligation was carried out at 12°C overnight. Transformation into bacteria was carried out using a heat shock protocol according to the manufacture's instructions. Potentially positive clones were identified by resistance to ampicillin (Amp) (100μg/ml), and X-gal colour selection. Positive colonies were selected and streaked in duplicate onto two selection plates.
Colonies from one of the selection plates were lifted (314) onto nitrocellulose filters (Sleicher & Schuell) and the filters probed with the appropriately labeled oligonucleotides to identify clones with the desired inserts (as described above). Positive colonies were selected and plasmid DNA was extracted (see DNA extraction).

Sequencing was performed using the double-stranded dideoxy sequencing method with the T7 Sequencing Kit (Pharmacia Fine Chemicals, Piscataway, NJ); both the universal and reverse primers were used. Briefly, the universal (or reverse) primer hybridizes to its complementary sequence located on the denatured TA plasmid vector, and the 3' hydroxyl group permits the synthesis of the complementary strand of DNA by T7 polymerase. Modified $\alpha^35$-dATP is incorporated into the newly synthesized DNA fragment. The elongating DNA reactions are subdivided into four separate "termination" reactions, each containing one of the four dideoxynucleotides. Incorporation of a dideoxynucleotide into DNA prevents further elongation (hence the name termination). The sequence of the DNA is deduced by electrophoresing the four reactions side-by-side.
CHAPTER III

HIGH FREQUENCY OF DETECTABLE NORMAL DJH REARRANGEMENTS IN SCID MICE

Contents of this chapter appeared in the *The Journal of Experimental Medicine, 1993, Volume 178: 1007-1016* by copyright permission of The Rockefeller University Press.

In the following chapter, DNA titration standards had been established by Dr. Yenhui Chang. Jakub Celler probed the indicated Southern blots and performed the densitometric analysis of the blots. Gillian Wu did the statistical analysis of the densitometric analysis. I performed PCR analysis (not shown) of BM DJH rearrangements and did all the sequencing and cloning.
INTRODUCTION

The formation of functional immunoglobulin (Ig) and T cell receptor (TCR) genes requires the rearrangement of several genetic elements encoding the variable regions: V, J and in some cases, D (19). The mechanism of this recombination is only now beginning to be elucidated, and most of what we know about the process is derived from studies of its substrates and its products (see (29, 61) for review). The process of rearrangement is mediated by an enzymatic system, the V(D)J recombinase, which is targeted by the recombination signal sequences (RSS) flanking the elements to be rearranged. RSS consist of three parts; a heptamer, a 12 bp or 23 bp spacer and a nonamer. As a result of V(D)J recombination, two types of junctions are formed: signal joints and coding joints (36). In the signal joints, which are usually precise, RSS are joined in a head to head orientation. Coding joints, which are usually imprecise, have a few nucleotides added and/or deleted at the coding termini.

V(D)J recombination is essential for normal lymphoid development. Mice deficient for the recombinase activating genes, RAG-1 and RAG-2, as well as SCID mice, have defective V(D)J recombinase activity (38, 58, 59, 237), lymphoid development in all these mice is arrested at the pro-B cell stage. The RAG proteins have been identified as the enzymes responsible for both RSS recognition, and subsequent cleavage (125, 128, 129, 139). Consequently RAG-deficient lymphoid cells do not attempt gene rearrangement (60, 132). Unlike the RAG-deficient mice, SCID lymphoid cells attempt gene rearrangement, but these attempts fail to produce functional receptors. This inability to undergo complete V(D)J recombination is a consequence of a defective double-strand DNA break (DSB) repair pathway (2-4). In Abelson murine leukemia virus (A-MuLV) transformed SCID bone marrow (BM) lines and in long term cultures of
SCID BM there are large deletions in the coding joints even though signal joint formation is relatively normal (38, 39, 199, 200, 236-238).

The SCID defect is leaky. Approximately 15% of all SCID mice develop detectable levels of oligoclonal serum Ig and T cell activity (7, 210-212). The incidence of leakiness varies with both the age of the mice and their environment (7). It has been postulated that a genetic reversion may occur in some lymphoid precursors permitting normal V(D)J recombination in subsequent daughter cells (45). In support of this hypothesis, some functional T cell clones derived from leaky SCID mice have normal rearrangements on their non-expressed TCR alleles (45). Moreover, a SCID A-MuLV line was found to have wild type V(D)J recombination activity (241). These observations are unexpected given the high frequency of abnormal coding joints found in A-MuLV SCID lines (39, 199, 236, 237). Normal TCR rearrangement at the δ DJ locus has been observed in SCID thymocytes (51, 322). However, there is no evidence of Vδ to DJδ joining or rearrangement at the β and γ loci of SCID thymocytes (243), a finding inconsistent with a genetic reversion being responsible for the observed leakiness.

Alternatively, SCID lymphoid cells may retain a higher level of V(D)J recombinase activity than predicted, so that a few B and T cells are produced each day depending on the chance occurrence of several rearrangement events. This hypothesis would explain the correlation of leakiness with both age and the cleanliness of the facility in which the mice are reared (1). Moreover, the SCID gene product has recently been shown to be the catalytic subunit of the DNA-dependent protein kinase (DNA-PK), a member of the phosphatidylinositol-3 kinase (PI-3 kinase) superfamily (8, 9, 305). SCID cells contain detectable DNA-PKcs protein, however at 1-10% the level observed in wild type cells (8, 9, 301). The SCID mutation has been identified as a premature ochre stop codon within the DNA-PKcs kinase domain that truncates the protein by 83 AA (301, 307).
conserved motifs required for catalysis in conventional PI-3-kinases are retained in the mutant DNA-PKcs, suggesting that the SCID kinase may be catalytically active (301, 307). No kinase activity has been ascribed to SCID DNA-PK as yet, however this may reflect the sensitivity of the assays used (301).

To gain more insight into SCID defect and its effect on the mechanism of V(D)J rearrangement, we have analyzed DJH recombination products in SCID mice. Using a quantitative PCR assay devised to determine the extent of DJH formation (26), we found DJH structures in SCID mice occur at a frequency only 10 to $10^2$ times less than in C.B-17 mice. Sequencing of these structures revealed that the SCID DJH repertoire is limited and resembles fetal liver DJH junctions, with few N insertions and predominant usage of reading frame (RF) 1. Moreover, 70% of these structures were potentially productive, indicating that normal V(D)J recombinants should be arising continually in SCID mice.
RESULTS

QUANTIFICATION OF DJH STRUCTURES

To quantitatively assay DJH structures by PCR we use primers flanking D and JH4 that results in a "ladder" of DJH1, DJH2, DJH3, and DJH4 when the products are analyzed by gel electrophoresis (Figure 2-1). This assay is designed such that aberrant rearrangements which juxtapose an upstream DH element within 1.2 kb of the JH4 primer site should be amplified (see Figure 2-1). Moreover, by simultaneously amplifying standards containing equimolar concentration of targets of each of the four DJH products, I can estimate their number in tissues. It was previously shown that none of these products are preferentially amplified by the DFS/JH4 primers (26). Figures 3-1A and 3-1B show examples of typical PCR amplifications of DJH rearrangements in BM and spleen DNA from individual SCID and C.B-17 mice. There are four discrete bands of sizes corresponding to DJH1, DJH2, DJH3, and DJH4 structures in all samples. The discrete nature of the bands was somewhat surprising due to the paucity and aberrant nature of the SCID coding junctions found previously (199, 237). I quantified these products using data derived from seven C.B-17 BM, five C.B-17 spleens, five SCID BM, four SCID spleens. One SCID mouse was serum Ig positive. The raw data from a typical quantification are shown in Table 3-1; a summary of the quantification data are shown in Table 3-2. The data are presented as the sum of the four JH's because in some cases individual SCID mice had unequal usage of the JH segments (Figure 3-1A, 3-1B and Discussion below). This unequal usage of JH elements was not observed in other strains (Figure 3-1A, 3-1B, and (26)). There were about $4.2 \pm 0.5 \times 10^5$ DJH products per SCID femur which is about 10% that of C.B-17 ($3.0 \pm 0.6 \times 10^6$). In the spleen, there were about $2.8 \pm 1.0 \times 10^5$ DJH structures in SCID and about $1.6 \pm 0.3 \times 10^7$ in C.B-17.
Figure 3-1: A. Southern blot analysis of PCR amplification of SCID and C.B-17+/+ bone marrow (BM) and spleen (SP) DNA using the DFS/JH4 primer pair. Amplifications of the titration standards were run in parallel.
B. Southern blot analysis of PCR amplified BM and SP DNA from non-leaky (Ig-) and leaky (Ig+) SCID and C.B-17+/+ mice. Titration standards were also amplified in parallel.
## TABLE 3-1
Typical Densitometric Analysis

<table>
<thead>
<tr>
<th>DNA Source</th>
<th>DJH1</th>
<th>DJH2</th>
<th>DJH3</th>
<th>DJH4</th>
<th>Total</th>
<th>Number per organ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CPM 1</td>
<td>Number 2</td>
<td>CPM</td>
<td>Number</td>
<td>CPM</td>
<td>Number</td>
</tr>
<tr>
<td>500 T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1000 T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5000 T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+/- BM -11</td>
<td>719</td>
<td></td>
<td>4816</td>
<td></td>
<td>1480</td>
<td></td>
</tr>
<tr>
<td>+/- Spleen -10</td>
<td>439</td>
<td></td>
<td>2790</td>
<td></td>
<td>1100</td>
<td></td>
</tr>
<tr>
<td>+/- Spleen -12</td>
<td>398</td>
<td></td>
<td>2490</td>
<td></td>
<td>1090</td>
<td></td>
</tr>
<tr>
<td>SCID BM -17</td>
<td>698</td>
<td></td>
<td>4662</td>
<td></td>
<td>273</td>
<td>&lt;500 (101)</td>
</tr>
<tr>
<td>SCID BM -19</td>
<td>274</td>
<td></td>
<td>1598</td>
<td></td>
<td>359</td>
<td>&lt;500 (385)</td>
</tr>
<tr>
<td>SCID BM -1</td>
<td>221</td>
<td></td>
<td>1215</td>
<td></td>
<td>521</td>
<td>919</td>
</tr>
<tr>
<td>SCID Spleen -18</td>
<td>123</td>
<td></td>
<td>506</td>
<td></td>
<td>154</td>
<td>&lt;500 (209)</td>
</tr>
<tr>
<td>SCID Spleen -20</td>
<td>146</td>
<td></td>
<td>672</td>
<td></td>
<td>104</td>
<td>&lt;500 (44)</td>
</tr>
</tbody>
</table>

1. output on the densitometer
2. number of structures calculated using MicroSoft Excel TRENDS
3. one µg is equivalent to 1.8 x 10^5 cells
4. calculation per organ is based on BM having 2 x 10^7 nucleated cells for all strains; C.B-17 +/- spleen having 1.5 x 10^8 nucleated cells; SCID spleen having 3 x 10^7 nucleated cells.
5. numbers identify DNA preparations
### TABLE 3.2

**Quantification of DJH Structures in BM<sup>1</sup> and Spleen<sup>2</sup>**

<table>
<thead>
<tr>
<th></th>
<th>C.B-17 +/+ BM</th>
<th>SCID BM</th>
<th>C.B-17 +/+ Spleen</th>
<th>SCID Spleen</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>6.7 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>3.0 x 10&lt;sup&gt;7&lt;/sup&gt;</td>
<td>2.1 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>1.6 x 10&lt;sup&gt;6a&lt;/sup&gt;</td>
<td>5.4 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>1.6 x 10&lt;sup&gt;7&lt;/sup&gt;</td>
<td>2.3 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>6.7 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>4.0 x 10&lt;sup&gt;5c&lt;/sup&gt;</td>
<td>1.3 x 10&lt;sup&gt;7&lt;/sup&gt;</td>
<td>5.7 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>4.3 x 10&lt;sup&gt;6b&lt;/sup&gt;</td>
<td>3.0 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>1.4 x 10&lt;sup&gt;7d&lt;/sup&gt;</td>
<td>1.2 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>1.9 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td>3.9 x 10&lt;sup&gt;5&lt;/sup&gt;</td>
<td>1.2 x 10&lt;sup&gt;7&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.3 x 10&lt;sup&gt;6&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Av. 3.1 x 10&lt;sup&gt;6&lt;/sup&gt; (SE) (± 0.72 x 10&lt;sup&gt;6&lt;/sup&gt;)</td>
<td>4.6 x 10&lt;sup&gt;5&lt;/sup&gt; (± 0.65 x 10&lt;sup&gt;5&lt;/sup&gt;)</td>
<td>1.7 x 10&lt;sup&gt;7&lt;/sup&gt; (± 0.33 x 10&lt;sup&gt;7&lt;/sup&gt;)</td>
<td>2.8 x 10&lt;sup&gt;5&lt;/sup&gt; (± 0.99 x 10&lt;sup&gt;5&lt;/sup&gt;)</td>
<td></td>
</tr>
</tbody>
</table>

1 2 x 10<sup>7</sup> nucleated BM cell for both strains per femur
2 1.5 x 10<sup>8</sup> nucleated spleen cells for C.B-17 +/+ and 3 x 10<sup>7</sup> spleen cells for SCID. 
<sup>a,b,c,d</sup> denote the average value of DJH rearrangements from a mouse sampled multiple times.

Each value was determined from a set of data analysed as detailed Table 3-1.

Mouse<sup>a</sup> was evaluated twice; 1.0 x 10<sup>6</sup> and 2.1 x 10<sup>6</sup> DJH rearrangements per femur were observed for this mouse.

Mouse<sup>b</sup> was also evaluated twice, and 3.9 x 10<sup>6</sup> and 4.7 x 10<sup>6</sup> DJH joins per femur were observed.

Mouse<sup>c</sup> was evaluated on three separate instances. 4.3 x 10<sup>5</sup>, 2.8 x 10<sup>5</sup>, and 4.7 x 10<sup>5</sup> DJH rearrangements per femur were found in these individual samplings.

Mouse<sup>d</sup> was evaluated twice and found to contain 1.6 x 10<sup>7</sup> and 1.1 x 10<sup>7</sup> DJH joins per spleen.
Serum Ig⁺ leaky SCID were not significantly different from non-leaky SCID. Although easily detectable in C.B-17 BM DNA, no DQ52JH products were detected upon amplification of SCID BM DNA using DQ52/JH4 primers. Since the sensitivity of the assay is about 40 targets per μg DNA (26), this result indicates there are less than 4.4 x 10³ DQ52JH rearrangements per femur (Southern analysis not shown). From these data, we conclude that SCID BM and spleen have correctly sized DJH1-4 products at about 10% and 1% respectively of the frequency found in C.B-17.

CHARACTERIZATION OF DJH STRUCTURES

Based on the lack of mature lymphoid cells in SCID mice and the aberrant nature of the coding junctions isolated from A-MuLV transformed lines from SCID mice, we had expected to see few, if any, discrete normal-sized DJH fragments. To search for more subtle differences in the SCID DJH joints, we cloned and sequenced the amplified products and compared them to those from C.B-17. Figure 3-2 shows the DJH structures derived from 6 and 27 week old BM of C.B-17 mice. Figure 3-3 shows the DJH structures from 6 week old SCID BM. Key features of the analysis are summarized in Tables 3-3 and 3-4.

DJH REARRANGEMENTS IN C.B-17 BONE MARROW

Fifty-seven DJH structures from C.B-17 were analyzed. Since no difference was observed between structures derived from 6 week and 27 week old mice, the data in Figure 3-2 are pooled from both sources. Deletions were present in most coding joints. An average of 9.1 nucleotides was deleted from D and/or the JH. N and P insertions which are a common feature in adult Ig gene rearrangement (83) were observed in 82% of the structures, with an average of 4.2 nucleotides
**Figure 3-2:** DNA sequences of C.B-17+/+ DJH structures analyzed. Clones designated by A and B represent separate amplifications of BM DNA from a single 6 week old C.B-17+/+ mouse. Clones designated with a C are derived from a 27 week old C.B-17+/+ mouse. A total of 57 DJH sequences were assessed, 56 were found to be unique. The number of nucleotides deleted from the germline sequence of the recombining elements is indicated. In some structures the deleted nucleotides could have come from either the DH or JH element, and are underlined. Possible P insertions and N additions are presented in upper and lower case lettering respectively under the appropriate columns. In frame stop codons are bolded and italicized. Possible point mutations or Taq induced errors are in lower case letters.
<table>
<thead>
<tr>
<th>Clone</th>
<th>In Frequency</th>
<th>$s$</th>
<th>Insertions</th>
<th>$s$</th>
<th>JH Frequency</th>
<th>JH Sequence</th>
<th>Total $s$</th>
<th>Total $d$</th>
<th>Insertion SF</th>
</tr>
</thead>
<tbody>
<tr>
<td>JHI B3</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>TGG TAC TTT GAT GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>JHI B4</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>TTT GAT GAT GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>JHI B5</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>JHI B6</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>JHI B7</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>JHI B8</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>JHI B9</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>JHI B10</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B11</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B12</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B13</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B14</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B15</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B16</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B17</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B18</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B19</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B20</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B21</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B22</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B23</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B24</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B25</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B26</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B27</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B28</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B29</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B30</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B31</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B32</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B33</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B34</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B35</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B36</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B37</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B38</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B39</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>JHI B40</td>
<td>TT TAT TAC TAG GUT AUT</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
<td>TTT TAT TAT TAG GUT TGG GGC CAA</td>
<td>DEL 1 1 1</td>
<td>11</td>
<td>0</td>
</tr>
</tbody>
</table>

Average # of Deletions for All Sequences: 3.813
Average # of Nucleotide Additions for All Sequences: 6.813
Average # of Nucleotide Additions for Only Sequences Containing Them: 6.813
being added. Nine structures (16%) contain possible P nucleotides, of these seven also have N insertions.

Although DH elements can be read in all three reading frames (RF) and in either orientation, RF usage is not random. In the C.B-17 BM, RFs 1, 2, and 3 were used in 44%, 18%, and 38% of the structures, respectively. RF1 is most markedly overused in the fetal repertoire where there are few if any N additions due to the lack of terminal deoxynucleotidyl transferase (TdT) activity (73). Terminal homologies of the recombining elements are thought to promote the RF1 usage in cases where there is no N addition in the absence of TdT (26, 82, 83, 87). Indeed, upon analysis of structures without N nucleotides (nt) (nine structures), the RF1 bias becomes apparent - 67%:11%:22% (Table 2-4).

Of the 56 C.B-17 structures, 35 (61%) were potentially productive; i.e. they lack stop codons and maintain the residues that constitute FR4 of the JH element (118). Rearrangements involving the JH3 element were the least productive (40%) because of a more frequent usage of RF3 which encodes many stop codons. The biased DH usage reported by us (26) and others (84, 190) was also present; genetic element Dfl16.1 was utilized in 81% of the DJH rearrangements.

**DJH REARRANGEMENTS IN SCID BONE MARROW**

DNA derived from the BM of three 6-week-old, Ig- SCID mice were individually amplified and cloned. Figure 3-3 shows the sequences of 57 DJH structures from SCID BM DNA, and key features are summarized in Tables 3-3 and 3-4.

The SCID DJH structures shown in Figure 3-3 are, in fact, quite similar to those of C.B-17 mice (Figure 3-2) and other strains (26, 82, 83). However, there are marked quantitative and qualitative differences, particularly in the degree of diversity. Of 57 SCID structures isolated from 4 individual mice, only 27 were
Figure 3-3: DNA sequences of the non-leaky 6 week old SCID BM DJH structures. A, B & C designate different SCID mice. Lower case Roman numerals represent separate PCR primary and secondary amplifications. A total of 57 SCID DJH joins were sequenced, only unique joins are presented here. Numbers in parentheses indicate the number of individual PCR amplifications in which a particular DJH join was isolated from the BM of the same mouse. The number of nucleotides deleted from the germline sequences of the recombining elements are indicated. In some cases deleted nucleotides may have come from either recombining element and are underlined. Possible P insertions and N additions are presented as upper and lower case lettering respectively under the appropriate columns. In frame stop codons are italicized and bolded. * indicates a DJH structure with a fusion of a partial JH2 to JH4. Mutations from germline sequences are indicated by lower case letters. Clones whose mutations are not shown in this chart have their clone name bolded.
<table>
<thead>
<tr>
<th>CLOSE</th>
<th>IN SEQUENCES</th>
<th>0 DELETED</th>
<th>INSERTIONS</th>
<th>0 DELETED</th>
<th>IN SEQUENCE</th>
<th>TOTAL 0 DELETED</th>
<th>0 INSERTIONS</th>
<th>RF</th>
</tr>
</thead>
<tbody>
<tr>
<td>JH1</td>
<td>TC TAT GAT GGT TAG TAC</td>
<td>0</td>
<td>4</td>
<td>TGG TAC TCC GAT GTC TGG GCC</td>
<td>Desp 2.9</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D1 1</td>
<td>TTT ATT ACT AGC GTA GTA G</td>
<td>4</td>
<td>tt</td>
<td>TGG TAC TCC GAT GTC TGG GCC</td>
<td>DFL 16.1</td>
<td>8</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>JH2</td>
<td>TC TAT GAT GGT TAG T</td>
<td>2</td>
<td>15</td>
<td>GG GCC CAA GCC</td>
<td>Desp 2.9</td>
<td>17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A1 1</td>
<td>TC TAT TAG G</td>
<td>8</td>
<td>3 (JH2)</td>
<td>AT TTT GAC TAC TGG GCC CAA GG</td>
<td>Desp 2.1-2.5</td>
<td>42</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A1 3</td>
<td>A1 1</td>
<td>3 (JH4)</td>
<td>A ACC TCA GCC</td>
<td>DFL 16.1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>A1 2</td>
<td>C TTA CTA TAG T</td>
<td>6</td>
<td>6</td>
<td>TAC TAC TAC GGG CA GC</td>
<td>Desp 2.1</td>
<td>12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B1 2</td>
<td>C TTA CTA TAG TAA CTA C</td>
<td>8</td>
<td>GT AG gatt</td>
<td>TAC TTT GAC TAC TGG GCC CAA GCC</td>
<td>Desp 2.1</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>JH3</td>
<td>T CTA CTA TAG TTA CTA C</td>
<td>0</td>
<td>cc cgg tt</td>
<td>T TAC TGG GCC</td>
<td>Desp 2.2</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A1 3</td>
<td>T CTA CTA TTA TTA CGA C</td>
<td>0</td>
<td>9</td>
<td>CT TAC TGG GCC</td>
<td>Desp 2.2</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C1 1(4)</td>
<td>T TTA CTA CGG TAG TAG CTA C</td>
<td>0</td>
<td>cc cgg tt</td>
<td>T TAC TGG GCC</td>
<td>DFL 16.1</td>
<td>10</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>JH4</td>
<td>TT TAT TAC TAC GGT AGT AGC TA</td>
<td>1</td>
<td>AT TAC TAT GCT ATG GAC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>A1 4</td>
<td>TT TAT TAC TAC GGT AGT AGC TA</td>
<td>1</td>
<td>AT TAC TAT GCT ATG GAC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A1 8</td>
<td>TT CAT TAC TAC GGC TA</td>
<td>1</td>
<td>AT TAC TAT GCT ATG GAC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>C1i 2</td>
<td>TT TAT TAC TAC GGT AGT AGC TAC</td>
<td>0</td>
<td>5</td>
<td>TAC TAT GCT ATG GAC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C1i 3</td>
<td>TT TAT TAC TAC GGT AGT AGC T</td>
<td>2</td>
<td>0</td>
<td>AT TAC TAT GCT ATG GAC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>C1i 4</td>
<td>TT TAT TAC TAC GGT AGT AGC TAC</td>
<td>1</td>
<td>1</td>
<td>AT TAC TAT GCT ATG GAC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C1i 5</td>
<td>TT CAT TAC TAC GGC TA</td>
<td>2</td>
<td>0</td>
<td>AT TAC TAT GCT ATG GAC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C1i 6</td>
<td>TT TAT TAC TAC GGC TAC</td>
<td>0</td>
<td>5</td>
<td>TAT GCT ATG AGC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>C1 2</td>
<td>TT CAT TAC TAC GGC TAC</td>
<td>4</td>
<td>4</td>
<td>AT TAC TAT GCT ATG GAC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>8</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>B1 1</td>
<td>CCT ATG AGT TTA AA</td>
<td>4</td>
<td>9</td>
<td>GCT ATG GAC TAC TGG GCT</td>
<td>Desp 2.1</td>
<td>13</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A1 5</td>
<td>T TTA CTA CTA CAG TAC TA</td>
<td>5</td>
<td>7</td>
<td>T CTA GGT ATG GAC TAC TGG GCT</td>
<td>DFL 16.1</td>
<td>12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A1 8</td>
<td>C TTA CTA TAC GT</td>
<td>5</td>
<td>4</td>
<td>C TTA GGT ATG GAC TAC TGG GCT</td>
<td>Desp 2.1</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Average # Deletion - all sequences:
- 28 ± 1
- 60 ± 49

Average # Deletion - unique sequences only:
- 25 ± 26
- 68 ± 73

Average # Nucleotide Addition - all sequences:
- 01 ± 03
- 12 ± 25
- 01 ± 03

Average # Nucleotide Addition - unique sequences only:
- 40 ± 52 ± 14 ± 07 ± 07
- 0 ± 17
- 0 ± 24
TABLE 3-3:

Comparison of DJH structures derived from SCID and C.B-17+/+ BM and BALB/c fetal liver (FL)

<table>
<thead>
<tr>
<th></th>
<th>SCID</th>
<th>C.B-17+/+</th>
<th>BALB/c FL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number Structures Analyzed</td>
<td>57</td>
<td>57</td>
<td>40</td>
</tr>
<tr>
<td>Average Deletions (Range)</td>
<td>8.8</td>
<td>9.1</td>
<td>6.7</td>
</tr>
<tr>
<td>Number Structures Containing Nucleotide Insertions</td>
<td>14</td>
<td>47</td>
<td>0</td>
</tr>
<tr>
<td>Average Number of Nucleotide Insertions for Above Structures</td>
<td>5.4</td>
<td>4.2</td>
<td>0</td>
</tr>
<tr>
<td>RF Usage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>63%</td>
<td>44%</td>
<td>70%</td>
</tr>
<tr>
<td>2</td>
<td>7%</td>
<td>18%</td>
<td>8%</td>
</tr>
<tr>
<td>3</td>
<td>30%</td>
<td>38%</td>
<td>22%</td>
</tr>
<tr>
<td>Potentially Productive Joins</td>
<td>70%</td>
<td>61%</td>
<td>90%</td>
</tr>
<tr>
<td>DFL 16.1 Usage</td>
<td>39%</td>
<td>81%</td>
<td>52%</td>
</tr>
</tbody>
</table>

Data is summarized from Figures 3-2 and 3-3. BALB/c FL data is derived from (26).
### TABLE 3-4:

Comparison of DJH Structures with and without N or P nucleotide additions

<table>
<thead>
<tr>
<th></th>
<th>Total Scid</th>
<th>Unique Scid</th>
<th>Total C.B-17 +/-</th>
<th>Unique C.B-17 +/-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Without N or P</td>
<td>Total</td>
<td>Without N or P</td>
</tr>
<tr>
<td>Number of Structures</td>
<td>57</td>
<td>43 (75%)</td>
<td>22</td>
<td>17 (77%)</td>
</tr>
<tr>
<td>RF 1</td>
<td>63%</td>
<td>79%</td>
<td>55%</td>
<td>59%</td>
</tr>
<tr>
<td>RF 2</td>
<td>7%</td>
<td>2%</td>
<td>9%</td>
<td>6%</td>
</tr>
<tr>
<td>RF 3</td>
<td>30%</td>
<td>19%</td>
<td>36%</td>
<td>35%</td>
</tr>
<tr>
<td>DFL 16.1</td>
<td>39%</td>
<td>21%</td>
<td>50%</td>
<td>32%</td>
</tr>
<tr>
<td>Usage</td>
<td>70%</td>
<td>88%</td>
<td>55%</td>
<td>71%</td>
</tr>
<tr>
<td>Potentially Productive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

distinct isolates and of these, only 22 were unique, compared to 56 out of 57 DJH rearrangements isolated from three individual C.B-17 mice. Moreover, only 14 of the 57 SCID structures (5 of the 22 unique ones) contained N or P nucleotide insertions. Of the five unique structures containing N or P, two contained N only, two contained possible P and N, and one structure contained possible P only. The average number of insertions for structures containing N and/or P was 5.4, (4.4 nt for unique structures) somewhat more than C.B-17 (4.2 nt). The mean number of nucleotides deleted from the recombined coding ends was 8.8, which is about the same as C.B-17 (9.1 nt) (Table 3-3).

Seventy percent (55% of the 22 unique structures) of the SCID structures could yield a functional Ig protein as defined by the lack of stop codons and the maintenance of FR4 residues in the JH elements. Dfl16.1 was overused, but less frequently than in C.B-17. Thirty-nine percent of the SCID (50% if the unique only are considered) structures use this DH element (Table 3-4). SCID DJH structures used RF1 more frequently than C.B-17. Of the 57 SCID DJH structures, 63%, 7% and 30% used RFs 1, 2 and 3 respectively (compared to 44%, 18% and 38% respectively for C.B-17 respectively). Interestingly, RF1 is used more frequently in both strains when structures joined without N and/or P insertions (Table 3-4). Moreover, 59% of the unique SCID joins involved joining by homology, a feature characteristic of fetal rearrangements (26, 82, 83). Taken together, these SCID DJH joins more closely resemble wild type fetal gene rearrangements with the comparable number of deleted nucleotides, the decreased incidence of N addition and concomitant increased incidence of both homology mediated joining and RF1 usage.
DISCUSSION

FUNCTIONAL DJH JOINTS IN SCID LYMPHOID PROGENITORS

In A-MuLV transformed lymphocytes from SCID mice, all DJH coding joints were grossly abnormal, containing large deletions of both DH and JH regions (236, 237). Therefore, we expected to observe few, if any, DJH joints in cells from SCID BM. To our surprise, only normal-sized DJH rearrangements were observed in the present study at 1-10% the frequency of wild type whole BM. Moreover, the frequency of progenitor cells containing potentially functional joints was also close to normal. We arrived at this latter conclusion by determining the frequency of normal DJH joints detected by PCR and the proportion of cells in BM belonging to the B cell lineage. In normal BM, approximately 30% of the cells belong to the B lineage. Of these cells, approximately two-thirds are in the pre-B cell stage (i.e., contain cytoplasmic μ) or B cell stage (i.e., express surface Ig). The remaining cells are in the pro-B stages. According to a study by Osmond (206), the early and intermediate pro-B cells in SCID mice are normal in frequency and proliferation kinetics. The late pre-B stage is markedly absent in SCID mice, so that overall the number of B lineage cells in the BM of SCID mice is only about 10% of the number found in the BM of normal mice. Since SCID mice contain approximately 10% as many DJH joints as normal mice (Table 3-2), many of the pro-B cells in SCID mice must contain potentially functional DJH joints.

These data appear to contradict previously published data on the frequency of normal DJH joints in pro-B cells transformed by A-MuLV (47, 236, 237) and those pro-B cells produced in long-term BM cultures (238, 323) derived from SCID mice. In both instances, few, if any, normal DJH joints were observed in cells from SCID mice. Several factors may contribute to this discrepancy: The
low number of B220+ cells in the BM of SCID mice (206) indicates that cells which fail to make a functional heavy chain gene are rapidly deleted in the BM. However, cells with nonfunctional and/or aberrant rearrangements may be rescued either by transformation with A-MuLV or by the growth conditions in long-term BM cultures. Thus, the abnormal rearrangements may be a result of the rescue of cells from programmed cell death, and they may not accurately reflect the ability of the recombinase system to carry out DJH recombination. Alternatively, the SCID cells rescued by transformation or culture conditions may be incapable of regulating either access and/or activity of the RAG proteins. This lack of control would permit continuous rearrangement of the Ig loci, which in conjunction with the DSB repair defect, would contribute to the gross deletions observed.

It is also possible that the major defect in gene rearrangement occurs when the cells attempt to form a VH-DJH joint. Quasi-normal DJH rearrangements have been reported in SCID thymocytes at the δ locus, while there was no evidence of Vδ rearrangement (243). Studies in long term bone marrow cultures and in A-MuLV transformed cells provided little evidence for attempted V to DJH recombination; thus it is possible that activation of this process is highly abnormal and leads to the gross deletions mentioned previously. Using primers which detect a large proportion of the VH joints, I investigated V-DJH recombination in SCID BM cells with a similar quantitative PCR assay of the type described in this paper. The resulting data is presented in Chapter IV.
FETAL-NATURE OF DJH JOINTS IN SCID MICE

Although the proportion of potentially functional DJH joints in SCID mice was similar to that observed for normal mice, there were unusual features of the joints in SCID mice. The incidence of N addition in the SCID DJH junctions is reduced compared to wild type (see Figures 3-2, 3 and Table 3-3). A similar low frequency of N addition is observed in fetal liver cells (Table 3-3, (26, 82, 83)), and is attributed to the low levels of TdT in fetal liver (73, 82). As reported by Osmond (206) the frequency of TdT+ cells in SCID BM is identical to that observed in the BM of normal mice. Thus the low level of N-nucleotides is unlikely to result from overall reduced TdT levels.

Approximately 60% of the SCID DJH junctions involved joining by homology (Figure 3-3). During fetal development, junctions routinely occur at regions of homology resulting in a severe restriction in the observed repertoire (26, 71, 82-84). This process of joining by homology is distinct from homologous recombination which involves an exchange of genetic information between a damaged DNA molecule and an undamaged partner (324). It is likely that the microregions of identity existing between the recombining elements help align the DNA ends and assist in their resolution in the presence of the SCID defect.

The third unusual feature of the DJH rearrangements in SCID mice was the restricted repertoire observed. All 56 DJH joins isolated from C.B-17 BM were unique. In contrast, exhaustive sequencing analysis revealed that of the 27 distinct DJH clones isolated from SCID BM, only 22 were unique. If, as discussed above, the SCID mutation results in a dissociation between the timing of gene rearrangement and the sequential transition of cells from pro-B cells to pre-B cells to B cells, it is possible that an abnormal expansion occurs in the late pro-B cell stage allowing minor clonal dominance of some DJH rearrangements. Support for this explanation can be seen in some of the Southern blots (Figures 3-
1B, 1C) where at times, one of the four DJH bands are unusually intense. Such intense bands may represent an expanded clone of a single cell that was able to resolve its coding ends. However, DJH rearrangements, with the exception of those in RF2, do not encode a protein product (83). Thus, it is difficult to envision how the successful resolution of these joints confers a proliferative advantage to the cells bearing them.

Covalently sealed coding end intermediates in V(D)J recombination were first detected in SCID thymocyte DNA (130). The initial inability to isolate such structures from normal lymphocytes lead to speculation that the SCID defect affects the resolution and/or generation of these structures (29, 130). Characterization of the SCID hairpin and RSS intermediates revealed them to be intact, involving no addition or deletion from the germline elements (132, 134). This observation coupled to the recent isolation of opened coding ends in wild type thymocytes (136), suggests that SCID product is involved in the resolution of the covalently sealed coding ends. The decrease in N additions in SCID junctions may suggest that TdT activity is compromised in the context of the SCID defect. Alternatively, joining of SCID coding ends may be facilitated by short stretches of identity between the two recombining elements. Thus, N addition to the coding ends may be detrimental to the resolution of the DNA ends in SCID lymphocytes.

**IMPLICATIONS FOR THE LEAKY PHENOTYPE**

As mentioned above, SCID is a leaky phenotype; normal immunoglobulin secreting lymphocytes arise in SCID mice (7). One model put forward to explain leakiness is the somatic reversion model, which proposes that a genetic reversion event in a SCID pro-B or earlier cell leads to clones of cells capable of normal gene rearrangement. Supporting evidence of the somatic reversion model comes
from the isolation of normal coding joints on the non-expressed alleles in T cells obtained from leaky SCID mice (45), and from the isolation of a SCID A-MuLV line with wild type recombinase activity (241). However, the normal-sized SCID DJH and TCR rearrangements reported are unlikely to have resulted from a reversion event. As a population, the SCID rearrangements are distinguished by a decrease in N addition and an increase in homology-mediated joining (data presented here, (46, 47, 242, 244)). SCID VDJH and TCR rearrangements also exhibit an increase in both the frequency and length of P addition (see Chapter IV, (46, 47, 242, 244)). If a reversion event were to have occurred, one would expect to observe no qualitative differences between the wild type and SCID rearrangements.
CHAPTER IV

NORMAL VDJH AND VJκ REARRANGEMENTS IN NON-LEAKY SCID MICE

Contents of this chapter have been submitted for publication under the title "Normal VDJH and VJκ rearrangements in non-leaky SCID mice" by Jacqueline L. Pennycook, Robert A. Phillips and Gillian E. Wu. I wish to thank D. Bouchard and S. Xie for their expertise in cell sorting. All other work was performed solely by the author.
INTRODUCTION

The critical event in B cell development is the production of a functional Ig surface receptor; consequently, early B cell development is inexorably linked to the rearrangement of the Ig genes. Elucidation of the precise mechanism(s) of V(D)J recombination is important for understanding both early B cell development and diseases, such as cancer, that can develop when rearrangement occurs inappropriately (325).

The process of V(D)J recombination involves juxtapositioning the variable (V), diversity (D) and joining (J) genetic elements to form the variable region exon of the Ig and T cell receptor (TCR) genes (see (28) for reviews). Conserved recombination signal sequences (RSS) that flank the elements to be rearranged, target the recombinase enzymatic complex. RSS consist of three components: a palindromic heptamer, a 12- or 23- bp spacer, and an AT-rich nonamer. Generally two types of junctions are formed during recombination: signal and coding joints. Signal junctions are characteristically precise, involving a direct ligation of RSS at their heptameric borders. Coding elements imprecisely join, typically exhibiting addition and deletion of a small number of nucleotides, in a process critical for generating receptor diversity.

Insights into the mechanism of gene rearrangement are provided by mutations which affect V(D)J recombination. SCID mice are characterized by a profound absence of mature lymphoid cells (see (1) for review), which is a consequence of an inability to efficiently complete V(D)J recombination. Although the frequency of signal joint formation is relatively unaffected by the scid defect, coding joint formation is reduced approximately 10^3-fold (38, 241). This finding suggests that the SCID mutation specifically impairs coding joint formation. Nonetheless, 50% of the SCID signal junctions exhibit abnormal nucleotide deletion and addition (38,
suggesting that either the initiation of recombination at the RSS, or the fidelity of signal end joining is influenced by the scid defect. Intact signal end intermediates were subsequently found in both SCID and wild type thymocytes, however only SCID thymocytes contain readily detectable levels of covalently sealed coding ends (88, 130, 132). Neither SCID coding nor signal end intermediates exhibit loss or addition of nucleotides from the germline sequences of the recombining elements (60, 132, 134), indicating that initiation of the recombination reaction, i.e., cleavage at the RSS, is unaffected by the SCID defect.

Clues into the nature of the SCID gene product came from studies showing that it is required for the repair of double strand (ds) DNA damage induced by irradiation and various chemotherapeutic agents (2-5). There is now strong evidence that the scid gene encodes the catalytic subunit of the DNA-dependent protein kinase (DNA-PK), a member of the phosphoinositidyl-3 (PI3) kinase superfamily (8, 9, 66, 303, 304). DNA-PKcs is part of an enzyme complex that is active at sites of double strand breaks (DSB) in DNA (286, 288). The precise mechanism of DNA-PK's activity is unknown; however, it has the capacity in vitro to phosphorylate a number of transcription factors, e.g. SP1 (285), RNA polymerase I (313), and p53 (11), a molecule involved in cell cycle control and apoptosis. DNA-PKcs protein is found in both cytosolic and nuclear extracts of wild type cells, however the tightly bound nuclear fraction of DNA-PKcs is not observed in SCID cells (301). Moreover, the level of SCID DNA-PKcs protein is reduced 10-10^2 fold compared to wild type (8-10, 301). SCID cells contain normal amounts of steady-state DNA-PKcs mRNA, indicating that the decline in DNA-PKcs protein is likely a consequence of protein instability (301, 307). Insight into the genetic nature of the SCID defect has been provided by the identification of a single base pair mutation that results in a premature stop codon within the C-terminal kinase domain (301, 307). This mutation is predicted to truncate the protein by 83 residues (301, 307),
leaving the conserved motifs required for kinase activity of conventional PI3 kinase family members intact (308). In vitro kinase assays failed to detect catalytic activity in the mutant protein, however this failure likely reflects the sensitivity of the assay used (301). The mutant DNA-PKcs can still bind DNA (301), suggesting that the mutation does not disrupt the association of the SCID DNA-PKcs with the Ku complex. Collectively, these data can account for the 15-30% of SCID mice which develop detectable serum Ig after 3 to 6 months of age (1). Indeed, we and others have found that bone marrow (BM) derived from non-leaky SCID mice contains a relatively high level of detectable normal DJH rearrangements (221, 245, 246). Moreover, other investigators have shown that under conditions that induce DNA repair, T cells make a high frequency of normal-sized rearrangements in the β chain locus (225-227), while the α locus seldom rearranges (228). The high frequency of DJH rearrangements observed in our earlier study seems inconsistent with the large deletions isolated from SCID A-MuLV lines and the very low frequency of normal B cells detected in SCID mice (1).

In this study, I have examined in detail the rearrangement of the IgH and κ light chain Ig loci in the BM of SCID mice. Normal VDJH and VJκ rearrangements are readily detectable, and sequence analysis reveals that many are productive. Several coding joints share overlaps in DNA sequence present in both recombining elements, suggesting that joining by homology may be responsible for rescuing coding joint formation. Moreover, this result suggests that joining by homology does not require functional DNA-PK.
RESULTS

VDJH REARRANGEMENTS IN SCID BONE MARROW

The VHALL primer recognizes at least 80% of all VH genes (316). Use of this primer with a primer 3' of the JH locus (JH4-3') permits detection of differently sized PCR products depending on whether VDJH rearrangement occurred at JH1, JH2, JH3 or JH4 (see Figure 2-2). The data shown in Figure 4-1 summarize typical PCR amplifications obtained with DNA from C.B-17 (Figure 4-1A) and SCID (Figure 4-1B, 1C) BM cells using these primers. To compare the VDJH rearrangement status of B lineage subfractions, I amplified DNA derived from C.B-17 B220+sIgM- and B220+sIgM+ BM cells, as well as SCID B220+sIgM- BM cells. The SCID mice used did not have any detectable B220+sIgM+ cells (data not shown).

As shown in Figure 4-1A, amplification of DNA from normal C.B-17 BM gives the expected four VDJH bands with the enriched fractions having a substantial increase in the number of rearrangements. Figures 4-1B and 4-1C show typical amplifications of VDJH rearrangements from different SCID mice at 6 weeks of age. For each mouse, three replicate samples from whole BM and the B220+sIgM- cells were amplified. All SCID mice contained detectable VDJH rearrangements, with VDJH4 being the most prominent.

In all experiments, we found that the SCID VDJH structures are comparable in size to those observed in C.B-17; there were no visible aberrant bands. As can be seen in Figure 4-1B and 4-1C, the JH usage varies not only among different SCID mice but also in different samplings of lysates derived from the same mouse. This variation likely indicates that the incidence of VDJH in some SCID mice is at the limits of detection for our PCR assay (as we had found previously for DJH rearrangements (26)). Indeed, studies by other laboratories generally have failed to detect SCID VDJH rearrangements, suggesting that the frequency
FIGURE 4-1: Southern blot analysis of VHALL/JH4 PCR amplification. (A) VDJH rearrangements in C.B-17 BM-derived cells (indicated +/+). The T cell line El-4 served as a non-target DNA control. (B) and (C) VDJH rearrangements in SCID whole BM (Wh BM) and B220+ cell populations respectively. Amplifications of three individual SCID mice, S1, S2, and S3, were performed in triplicate. For comparison, (B) and (C) also show amplification of the equivalent population of C.B-17 mice. 10⁴ cell equivalents were amplified for all samples except C.B-17 B220+IgM+ BM population, where 10³ cell equivalents were amplified. The B220+ enriched fractions were purified using the FACStar cell sorter (see Chapter II) and all samples are >95% pure. Amplification for the non-rearranging α-actin gene served as an internal control for DNA loading.
Figure 4-2: Southern blot analysis of VDJH amplification of C.B-17 and SCID whole BM (Wh BM) DNA. CB-35 and CB-45 represent two wild type A-MuLV cell lines in a VDJH/VDJH configuration. Numbers indicate DNA derived from different SCID mice.
TABLE 4-1

Typical Phosphorimager Analysis of VDJH Rearrangements Observed in C.B-17 and SCID Mice Whole and Enriched BM Cells*.

<table>
<thead>
<tr>
<th>Population</th>
<th>VDJH1 (x/+)</th>
<th>VDJH2 (x/+)</th>
<th>VDJH3 (x/+)</th>
<th>VDJH4 (x/+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C.B-17 Wh BM</td>
<td>160 3.2</td>
<td>510 3.6</td>
<td>890 3.6</td>
<td>1600 3.6</td>
</tr>
<tr>
<td>C.B-17 B220* IgM^-</td>
<td>2100 1.6</td>
<td>5400 1.5</td>
<td>10000 1.6</td>
<td>16000 1.5</td>
</tr>
<tr>
<td>C.B-17 B220* IgM^+</td>
<td>1200 1.6</td>
<td>3400 1.5</td>
<td>5000 1.4</td>
<td>8500 1.5</td>
</tr>
<tr>
<td>SCID Wh BM 1</td>
<td>0.049 34</td>
<td>0.79 3.0</td>
<td>2.1 8.2</td>
<td>42 3.1</td>
</tr>
<tr>
<td>SCID Wh BM 2</td>
<td>0.73 1.1</td>
<td>0.043 63</td>
<td>5.1 1.2</td>
<td>4.8 2.7</td>
</tr>
<tr>
<td>SCID Wh BM 3</td>
<td>N.D.</td>
<td>0.005 7.4</td>
<td>N.D.</td>
<td>2.6 1.3</td>
</tr>
<tr>
<td>SCID B220^+ 1</td>
<td>0.61 190</td>
<td>17 5.4</td>
<td>24 1.7</td>
<td>74 1.9</td>
</tr>
<tr>
<td>SCID B220^+ 2^†</td>
<td>2.0 2.7</td>
<td>9.3 2.3</td>
<td>1.6 6.2</td>
<td>23 2.2</td>
</tr>
<tr>
<td>SCID B220^+ 3</td>
<td>0.009 22</td>
<td>3.7 1.4</td>
<td>0.33 110</td>
<td>42 1.5</td>
</tr>
</tbody>
</table>

*Geomeans and SD for indicated BM populations are provided. The above values represent signals above background levels (see CHAPTER II Quantification) for the experiment shown in Figure 4-1, normalized for the non-rearranging a-actin gene.
† C.B-17 B220^+ IgM^- had 10^3 rather than 10^4 cell equivalents amplified to prevent saturation of the PCR assay. N.D. denotes samples where the level of VDJH rearrangement is not detectable above background.
‡ Values presented for S2 B220^+ fraction are derived from only two lanes, as the third lane was a failed PCR reaction (data not shown).
### TABLE 4-2

Relative Frequency of SCID VDJH Rearrangement as Compared to C.B-17 B lineage cells*.

<table>
<thead>
<tr>
<th></th>
<th>VDJH1</th>
<th>VDJH2</th>
<th>VDJH3</th>
<th>VDJH4</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCID Wh BM 1</td>
<td>0.04%</td>
<td>0.15%</td>
<td>0.24%</td>
<td>2.6%</td>
</tr>
<tr>
<td>SCID Wh BM 2</td>
<td>0.46%</td>
<td>0.01%</td>
<td>0.57%</td>
<td>0.30%</td>
</tr>
<tr>
<td>SCID Wh BM 3</td>
<td>N.D.</td>
<td>&lt;0.01%</td>
<td>&lt;0.01%</td>
<td>0.16%</td>
</tr>
<tr>
<td>SCID B220+ 1</td>
<td>0.03%</td>
<td>0.33%</td>
<td>0.25%</td>
<td>0.50%</td>
</tr>
<tr>
<td>SCID B220+ 2</td>
<td>0.10%</td>
<td>0.17%</td>
<td>0.02%</td>
<td>0.14%</td>
</tr>
<tr>
<td>SCID B220+ 3</td>
<td>&lt;0.01%</td>
<td>0.07%</td>
<td>&lt;0.01%</td>
<td>0.26%</td>
</tr>
</tbody>
</table>

* The above frequencies were derived by comparing the SCID geomeans derived from the phosphorimaging analysis of the indicated rearrangements to those of the corresponding wild type VDJH rearrangements, shown in Table 4-1. Similar populations were compared in the derivation of the above frequencies, i.e. the frequency of VDJH rearrangement in SCID Wh BM was compared to that of C.B-17 Wh BM.
of SCID rearrangements is below the limit of sensitivity for the assays used (60, 213, 221, 299).

Occasionally, some SCID mice contain high levels of a single rearrangement. For example, in one SCID mouse a VDJH\text{4} rearrangement equivalent to wild-type levels was present (Figure 4-2, mouse C in Figure 4-3), while the level of rearrangements to the other JH elements were comparable to the low SCID levels shown in Figure 4-1B and 4-1C. Perhaps in these cases there was an unusual expansion of a cell bearing this rearrangement (see Discussion).

To obtain a relative quantification of the rearrangements of each JH element in the different strains of mice, the intensity of the bands were quantified with a PhosphorImager (Table 4-1). Table 4-2 compares the level of SCID VDJH rearrangement to C.B-17. Although we isolated analogous B lineage populations by purifying B220\textsuperscript{+}sIgM\textsuperscript{-} cells, inherent developmental differences exist between the two strains of mice: 48% of the C.B-17 B220\textsuperscript{+}sIg\textsuperscript{-} fraction are \textit{cu}\textsuperscript{+}, whereas \textit{cu}\textsuperscript{+} cells are undetectable in the SCID B220\textsuperscript{+}sIg\textsuperscript{-} population (206, 207). The values presented in Table 4-2 show an estimate of the relative frequency of SCID VDJH rearrangement in the given populations compared to C.B-17 without regard for the \textit{cu} status of the cell.

**High Frequency of Productive VDJH Rearrangement in SCID Mice**

The VDJH PCR products were subcloned and sequenced as described. Figure 4-3 shows the sequence of 21 unique clones derived from multiple primary amplifications of DNA from three different SCID mice. A summary of the number of deletions and the incidence of N/P addition at the VDH and DJH borders is also provided in Figure 4-3.

These data have several interesting features: First, out of 21 unique clones, 18 are in frame (86%), only clones A9, B1 and B9 are non-productively rearranged. VDJH joins derived from wild type \textit{cu}\textsuperscript{-} pro-B/pre-B\textsuperscript{I} cells and \textit{cu}\textsuperscript{+}
pre-BII cells are 25-40%, and 80% in frame respectively (84, 192). The frequency of productive SCID VDJH rearrangements is comparable to that observed in normal pre-BII cells suggesting the SCID B220+ cells bearing these rearrangements have survived selection. None of these structures had larger than normal deletions at either the V-DH or D-JH borders, a circumstance consistent with selection. Moreover, many more deletions are observed between the JH and DH elements than between the VH and DH elements in both SCID and wild type rearrangements.

• Many SCID VDJH Arise By Homology-Mediated Recombination

Clones A2 and A4 contain sequences at their respective V-DH borders that may have been derived from either the germline VH or DH recombining elements. These regions of identity have been arbitrarily assigned to the VH element in Figure 4-3, and inadvertently give the impression that these clones arose from a direct VH to JH joining event. However, it is more likely that these junctions arose from homology-mediated recombination between the sequences of identity shared by the recombining VH and DH elements. Interestingly, the regions of identity in the A2 and A4 joins encode the invariant Cysteine (Cys) residue found at codon 92 of all VH elements that is critical for Ig folding (118). Seven other clones have at least two base pairs that may have been derived from either of the recombining elements or from their P additions. Thus, at least 40% (9/21) of the VDJH joins may have arisen by homology mediated recombination involving two to seven nucleotides of identity. The incidence of homology mediated joining in SCID IgH rearrangements is much higher than the 6% and 12% observed at the V-DH and D-JH borders respectively of wild type adult IgH rearrangements (84).

• Imprecise RSS Cleavage in SCID VDJH Structures

Clones A24 and B7 in Figure 4-3 lack detectable DH elements. These joins appear to be derived from recombination events where cleavage occurred on the
FIGURE 4-3: SCID VDJH rearrangements derived from 6-week old SCID mice BM. Upper case letters indicate clones derived from an individual mouse. P insertions are presented in upper case lettering under the appropriate column, while N additions are in lower case lettering. Sequences that can be attributed to either recombining element or the P additions derived from these elements are underlined. Isolates designated with * possess identical DJH borders, as do isolates designated with †. Clones A16 and A17 differ by one base pair at codon 88 (sequence not presented above), resulting in Phe (TTT) and Ser (TCT) residues respectively. Sequences differences within the VH primer binding site are not considered relevant given the degeneracy of the primer(s) in question. ‡ indicates clones exhibiting cleavage on the wrong side of the RSS heptamer, consequently these clones do not have a DH element designation. The germline heptameric sequences are listed in italicized lower case under the DH element column. @The average number of deletions from the germline elements is calculated in three ways. (1) All homologies existing between VH and DH elements are arbitrarily assigned to the VH element; sequences which consequently appear to have deleted the DH element have the number of deletions assigned to the 5' end of the DH segment. All homologies between the DH and JH element are assigned to the JH element; thereby maximizing the number of deletions to the 5' end of the DH segment. (2) All homologies between VH and DH elements are assigned to the DH element. All homologies between the DH and JH element are assigned to the JH element, thus maximizing the number of deletions to the 3' ends of the VH and DH elements. (3) Number of deletions observed from recombining elements of sequences that exhibit distinct and recognizable VH, DH and JH elements, (i.e. clones A2, A4, A24 and B7 are excluded from the calculation). All homologies between the DH and JH element are assigned to the 5' end of the JH element, thereby maximizing the number of deletions to the 3' end of the DH segment.
<table>
<thead>
<tr>
<th>CLONE</th>
<th>VH ELEMENT</th>
<th># INSERTIONS</th>
<th># DELETED</th>
<th># INSERTIONS</th>
<th>DELETED</th>
<th>JS ELEMENT</th>
<th># INSERTIONS</th>
<th># DELETED</th>
<th>JS ELEMENT</th>
<th>DELETED</th>
<th>VH</th>
<th>DELETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>TAC TAC TGT AGA</td>
<td>0</td>
<td>0</td>
<td>ZCE ACT ACG</td>
<td>7</td>
<td>0</td>
<td>CAC TAC TAC</td>
<td>11</td>
<td>GAT UTC TOG</td>
<td>102</td>
<td>2.1</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>TAC TAC T</td>
<td>6/15</td>
<td>21/17</td>
<td>(14/8)</td>
<td>0</td>
<td>GAT UTC</td>
<td>0</td>
<td>C TAC TAC TAC TUV GAT UTC TOG</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>TAC TAC TGT GCA AAA</td>
<td>0</td>
<td>a</td>
<td>GT ACG</td>
<td>0</td>
<td>GAT UTC</td>
<td>52</td>
<td>16.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>TAC TAC T</td>
<td>6/15</td>
<td>21/17</td>
<td>(14/8)</td>
<td>0</td>
<td>GAT UTC</td>
<td>0</td>
<td>C TAC TAC TAC TUV GAT UTC TOG</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>TAC TAC TGT GCA AAA AAT</td>
<td>0</td>
<td>A</td>
<td>GT ACG</td>
<td>0</td>
<td>GAT TAC</td>
<td>52</td>
<td>16.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>TAC TAC TGT GCA AGA AA</td>
<td>0</td>
<td>TAC TACT</td>
<td>a a</td>
<td>0</td>
<td>GAT UTC</td>
<td>0</td>
<td>C TAC TAC TAC</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B4</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B5</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B6</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>TAC TAC TGT GCA</td>
<td>5</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A9</td>
<td>TAC TAC TGT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A10</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A16</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A17</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A18</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A21</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A24</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A25</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A27</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>TAC TAC TGT GCA AGA</td>
<td>0</td>
<td>TAC TACT</td>
<td>52</td>
<td>16.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Average # of Insertions:**
- All Sequences: 1.3 ± 2.4
- Only Sequences: 1.5 ± 2.9
- With Np: 1.6 ± 2.5

**Average # of Deletions:**
- All Sequences: 1.3 ± 2.4
- Only Sequences: 1.5 ± 2.9
- With Np: 1.6 ± 2.5

**Average # of N insertions:**
- All Sequences: 2.2 ± 4.4
- Only Sequences: 2.2 ± 4.4

**Average # of P insertions:**
- All Sequences: 2.5 ± 2.5
- Only Sequences: 2.5 ± 2.5

**Average # of JS insertions:**
- All Sequences: 2.5 ± 2.5
- Only Sequences: 2.5 ± 2.5

**Average # of VH insertions:**
- All Sequences: 2.5 ± 2.5
- Only Sequences: 2.5 ± 2.5
side of the heptamer of the JH element. Most of the heptamer sequences normally found adjacent to the germline JH elements are still present in both of the VDJH structures.

- **SCID Mice Exhibit Increased Incidence of P addition & Reduced N Insertions**

In normal mice, P insertions occur in 5-14% of VDJH junctions and 95% of these junctions have only two P nucleotides inserted (42, 43). The frequency of P addition in SCID VDJH joins is much higher, 67% of the clones have P insertions, 50% of which ≥ three nucleotides in length (see Figure 4-3).

In wild-type mice, most (>95%) VDJH structures from BM contain N additions (83, 84). In SCID mice, the incidence of N addition is 52% at the D-JH border with an average of 2.8 ± 1.7 bases added per join (see Figure 4-3). 43% of the V-DH borders contain N addition, averaging 5.2 ± 6.4 nucleotides added to these junctions. 67% percent of the 21 unique structures contain N addition at one or both of their junctions. Thus, the incidence of N addition is considerably reduced compared to normal mice.

The data shown in Figure 4-3 contain a unique difference from the data presented previously for DJH rearrangements in SCID BM ((245), see Chapter II). In that study, I observed very few insertions in the DJH junction: Only five of the 22 unique DJH sequences (23%) contained N or P additions, compared with 81% of the unique VDJH rearrangements presented here. However, independent analysis of the VDH and DJH borders within the SCID VDJH rearrangements, reveals that only 12/21 (57%) of the unique VDJH joins have nucleotide additions at their D-JH border, and four of these 12 VDJH joins share an identical DJH junction. Since the DJH PCR assay does not detect DJH joins within a VDJH rearrangement, and because repeated isolation of the same DJH join was only considered once in the DJH analysis, these unique VDJH structures which have identical D-JH borders
make the frequency of nucleotide insertion in VDJH rearrangements appear artificially high.

SCID MICE UNDERGO "NORMAL" & ABERRANT \( \kappa \) REARRANGEMENT

I used \( \kappa \)-specific PCR primers to examine rearrangements at the \( \kappa \) light chain locus in DNA from normal and SCID BM cells. Amplification of DNA from normal BM gives the expected four bands (Figure 4-4A), and, similar to the VDJH gene rearrangement amplifications, the highest incidence of \( \kappa \) rearrangement was observed in the B220+slgM+ population.

Figure 4-4B shows the result of amplifying DNA from cells in SCID mice. In both the whole BM and the B220 enriched populations, one can detect rearrangements at the \( \kappa \) locus giving the expected sized bands for VJ\( \kappa \)1, 2, 4 and 5. However, there are two abnormally sized bands migrating slightly faster than VJ\( \kappa \)4 and VJ\( \kappa \)5 (referred to as \( \Delta \)VJ\( \kappa \)4 and \( \Delta \)VJ\( \kappa \)5 respectively). These aberrant bands were detected in both whole BM and the B220+ enriched population from SCID donors, but were not detected in any population from normal C.B-17 mice.

Sequencing analysis of these aberrant-sized PCR products determine that the junctions involve five to seven bases of identity between the V\( \kappa \) and J\( \kappa \) elements (see Figure 4-6). However, the region of overlap within the V\( \kappa \) element involves the terminal five to seven nucleotides of the site to which the V\( \kappa \) con primer hybridizes. In order to determine whether the observed aberrant structures represent true products of recombination or a PCR-mediated event, I designed the V\( \kappa \) con-S primer which lacks the region of identity shared between the V\( \kappa \) con primer and the J\( \kappa \)4 and J\( \kappa \)5 elements (see Figure 2-3 and Table 2-1). Figure 4-5 shows the results of parallel amplifications of BM-derived lysates from both wild type and SCID under identical PCR cycling conditions using the V\( \kappa \) con and V\( \kappa \) con-S primers. Panel A shows amplification using V\( \kappa \) con/J\( \kappa \)5 with
FIGURE 4-4: Southern blot analysis of κ rearrangement in C.B-17 and SCID mice. (A) C.B-17 BM-derived cells (indicated +/+). The T cell line El-4 was used as a non-target DNA control. (B) κ rearrangements in SCID whole BM and B220+ cells from three individual 6 week old SCID mice. Aberrantly-sized PCR products migrating faster than VJκ4 and VJκ5 are referred to as ΔVJκ4 and ΔVJκ5, respectively. All B220+ fractions were purified using a FACStar Cell Sorter, and are >95% pure.
FIGURE 4-5: Southern blot analysis of κ rearrangement in SCID mice using (A) Vκcon and (B) Vκcon-S primers. DNA samples were subjected to identical PCR cycling conditions. Aberrant sized PCR products are labeled as described in Figure 4-3. B220+ cells were isolated by FACS cell sorting and are >95% pure. Lysates from C.B-17 spleen were amplified at 10⁴ cells/reaction, and the amplified product was then serially diluted in order to demonstrate that the aberrant sized bands are not present in wild-type cells.
**TABLE 4-3**

Typical Phosphorimager Analysis of VJκ Rearrangements in SCID and C.B-17 Mice*.

<table>
<thead>
<tr>
<th></th>
<th>Arbitrary Phosphorimager Value x 10^-3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VJK 1</td>
</tr>
<tr>
<td>C.B-17 Wh BM</td>
<td>5.9 x/+ 12</td>
</tr>
<tr>
<td>C.B-17 B220^+IgM^-</td>
<td>200</td>
</tr>
<tr>
<td>C.B-17 B220^+IgM^-</td>
<td>644</td>
</tr>
<tr>
<td>SCID Wh BM 1</td>
<td>N.D.</td>
</tr>
<tr>
<td>SCID Wh BM 2</td>
<td>0.020</td>
</tr>
<tr>
<td>SCID Wh BM 3</td>
<td>0.043</td>
</tr>
<tr>
<td>SCID B220^+ 1</td>
<td>0.60</td>
</tr>
<tr>
<td>SCID B220^+ 2</td>
<td>N.D.</td>
</tr>
<tr>
<td>SCID B220^+ 3</td>
<td>0.15</td>
</tr>
</tbody>
</table>

*The above data is derived from the experiment shown in Figure 4-4. All data above are the geometrical means of duplicate reactions, with the exception of C.B-17 whole BM which was performed in triplicate. In all samples 10^4 cells were amplified. The SD is provided for only C.B-17 whole BM population performed in triplicate.
TABLE 4-4

Comparison of VJk rearrangements in SCID whole and enriched BM to wild-type C.B-17*.

<table>
<thead>
<tr>
<th>SCID Wh BM 1</th>
<th>VJK 1</th>
<th>VJK 2</th>
<th>VJK 4</th>
<th>VJK 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>N.D.</td>
<td>0.7%</td>
<td>0.04%</td>
<td>0.5%</td>
<td></td>
</tr>
<tr>
<td>SCID Wh BM 2</td>
<td>0.3%</td>
<td>0.1%</td>
<td>2.0%</td>
<td>1.0%</td>
</tr>
<tr>
<td>SCID Wh BM 3</td>
<td>0.7%</td>
<td>0.02%</td>
<td>0.4%</td>
<td>0.8%</td>
</tr>
</tbody>
</table>

| SCID B220+ 1 | 0.3%  | 0.01% | <0.01% | 0.1%  |
| SCID B220+ 2 | 0.1%  | 0.03% | 0.3%   | 0.08% |
| SCID B220+ 3 | N.D.  | N.D.  | 0.01%  | 1.2%  |

*The above comparison is derived from the geomean data presented in Table 4-3.
N.D. indicates samples in which the level of VJk rearrangement was not detectable above background.
wild-type spleen DNA as a control. Panel B shows amplification of the same samples using the Vkcon-S primer. Aberrant and wild-type κ rearrangements are readily detectable under both primer conditions, indicating that the aberrant structures represent true products of rearrangement and are unlikely to be PCR artifacts. The apparent reduction in frequency of SCID κ rearrangement is most reflects the lower Tm of Vkcon-S, consequently amplification was not as efficient as with Vkcon. No experiments showed aberrant bands in the DNA derived from wild-type mice, with either Vkcon or Vkcon-S primer ((318), data presented here).

Table 4-3 summarizes the quantification (using the PhosphorImager) of the rearrangements detected in the amplification shown in Figure 4-4. Table 4-4 shows the relative frequency of Igκ rearrangements in SCID as a percentage of those detected in normal BM. This comparison indicates that SCID BM contains detectable non-aberrant VJκ rearrangements in the order of 0.01-1% of normal B220+sIg- cells. As discussed above, the relative frequency of κ rearrangements is likely an underestimation because of differences in the frequency of moe mature B lineage cells in the two strains (206). Moreover, the degree of underestimation of the frequency of κ rearrangement is further compounded by the fact that IgL rearrangement is enhanced by the signaling through the pBCR (78, 94, 326).

**N Insertion in SCID κ Rearrangements**

SCID κ rearrangements were cloned and sequenced from multiple primary and secondary PCR’s using either of the two Vkcon primers. The junctions are shown in Figure 4-6. Thirty unique sequences were isolated from seven different SCID mice; 25 sequences are derived from the normal-sized VJκ rearrangements, 44% of which are in frame. These structures exhibit comparable deletions from the Vk and Jκ elements, 1.4 ± 2.5 and 2.6 ± 2.4 bases respectively. 96% of wild type VJκ rearrangements contain <3 base pairs in total deleted from the κ recombining
FIGURE 4-6: Sequence analysis of SCID κ rearrangements. Letters indicate clones derived from individual mice. Clones given a VS designation indicate PCR amplification with Vκcon-S primer. Δ denotes aberrant-sized rearrangements derived from ΔVJκ4 or ΔVJκ5. Underscored sequences represent those sequences that may have been derived from either recombining element. Sequences overlapping with the Vκcon or Vκcon-S primer sites are italicized. P insertions are indicated by upper case lettering, while N insertions are in lower case. Brackets denote variations in sequences containing the same VJκ border. These sequence differences correspond to sites of degeneracy within the Vκcon and Vκcon-S primers (see Table 2-1), consequently clones bearing them are not considered to be different.
<table>
<thead>
<tr>
<th>Clone</th>
<th>Vx Element</th>
<th># INSERTION Deleted</th>
<th>In P N P Deleted</th>
<th>Jx Sequence</th>
<th>Jx FRAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>ACT TAT TAC TGT CAA CAG TAT C</td>
<td>9</td>
<td>C A C</td>
<td>0</td>
<td>GT GGA GGT TCG GGT GAG</td>
</tr>
<tr>
<td>B5</td>
<td>GTT TAT TAC TGT CAG CAA TAT TAT AGC TAT CTT C</td>
<td>0</td>
<td></td>
<td>2</td>
<td>GG ACG TTC GGT GGA</td>
</tr>
<tr>
<td>B6</td>
<td>GTT TAT TAC TGT TGT CAG AAT GAT CAT AGT TAT CTT C</td>
<td>0</td>
<td>tc</td>
<td>3</td>
<td>GG ACG TTC GGT GGA</td>
</tr>
<tr>
<td>B7</td>
<td>ACC TAT TAC TGT CAG CAA GCA ACA TTA TAG TAC TCC</td>
<td>0</td>
<td></td>
<td>2</td>
<td>G ACG TTC GGT GGA</td>
</tr>
<tr>
<td>G2</td>
<td>ACT TAT TAC TGT TGT CAA TAT TTT TGG GGT ACT CCT C</td>
<td>0</td>
<td></td>
<td>2</td>
<td>G ACG TTC GGT GGA</td>
</tr>
<tr>
<td>G3</td>
<td>GTT TAT TAT TGC TGG CAA GCT GCT ACT CACT T</td>
<td>7</td>
<td></td>
<td>1</td>
<td>TG GAC GGT CAC TCG AGG</td>
</tr>
<tr>
<td>G4</td>
<td>AGT TAT TAT TGC TGG CAA GGT ACA CAT T</td>
<td>6</td>
<td></td>
<td>1</td>
<td>TG GAC GGT CAC TCG AGG</td>
</tr>
<tr>
<td>G5</td>
<td>GTT TAT TAC TGT CAG CAA CAT TAT AGT ACT CCT CC</td>
<td>0</td>
<td></td>
<td>3</td>
<td>G ACG TTC GGT GGA</td>
</tr>
<tr>
<td>B8</td>
<td>ACT TAT TAC TGC CAG TGG AGT GGT TAC CCA TGC</td>
<td>0</td>
<td>t c</td>
<td>4</td>
<td>CAC GGT CAC GAG GGG</td>
</tr>
<tr>
<td>B9</td>
<td>GTT TAT TAC TGT GCT CAA AAT CCA GTA CC</td>
<td>0</td>
<td></td>
<td>2</td>
<td>G ACG TTC GGT GGA</td>
</tr>
<tr>
<td>B10</td>
<td>ACC TAT TAC TGT CAG CAA AAT GAT GAT GC</td>
<td>0</td>
<td></td>
<td>1</td>
<td>G TAC ACG TTC GGA GGG</td>
</tr>
<tr>
<td>B11</td>
<td>AGA TAT TAC TGT CAA CAT TAT TAT GAT CCT</td>
<td>2</td>
<td></td>
<td>4</td>
<td>G ACG TTC GGT GGA</td>
</tr>
<tr>
<td>B12</td>
<td>GTT TAT TAT TGC TGG CAA CTA GCA GTA CCT CC</td>
<td>0</td>
<td></td>
<td>2</td>
<td>TA CAC GGT CAC GAG GGG</td>
</tr>
<tr>
<td>B13</td>
<td>GGC ACT TAT TTT TGG CAA CAG GGT AAT AGC CCT CC</td>
<td>0</td>
<td></td>
<td>1</td>
<td>G TAC ACG TTC GGA GGG</td>
</tr>
<tr>
<td>B14</td>
<td>GTT TAT TAC TGT TGG CAA GGT ACA CAT TTT CCT CAT</td>
<td>0</td>
<td></td>
<td>4</td>
<td>CAC GGT CAC GAG GGG</td>
</tr>
<tr>
<td>B15</td>
<td>GTT TAT TAT TGC CGG CAA GGT ACA CAT TTT CCT CC</td>
<td>0</td>
<td>ta cac</td>
<td>11</td>
<td>GGA GGA</td>
</tr>
<tr>
<td>B16</td>
<td>GTT TAT TAT TGC CAG CAA GGT ACA TTT CCT TC</td>
<td>0</td>
<td>t cac c</td>
<td>6</td>
<td>GCT GGT CAA GGA</td>
</tr>
<tr>
<td>B17</td>
<td>AGA TAT TAC TGT CAA CAT TAT TGC TAC TCC TC</td>
<td>0</td>
<td></td>
<td>4</td>
<td>C ACG TTC GGA GGG</td>
</tr>
<tr>
<td>ΔH1</td>
<td>GCC CTC AGG TTC AGT GGC TGG GG(A/G) TCT</td>
<td>88/81</td>
<td>17-21</td>
<td>GGA/G CAC AAG TGG AAA TAA</td>
<td>4</td>
</tr>
<tr>
<td>V3 A1</td>
<td>AGT TAT TAC TGC CAG CAG TGG AGT AGT AAC CCA C</td>
<td>1</td>
<td></td>
<td>0</td>
<td>GC TCA CGT TCG GGT CTG</td>
</tr>
<tr>
<td>V3 A2</td>
<td>ACT TAT TAC TGC CAG CAG TGG AGT AGT AAC CCA C</td>
<td>0</td>
<td></td>
<td>0</td>
<td>GC TCA CGT TCG GGT CTG</td>
</tr>
<tr>
<td>V3 A3</td>
<td>GAT TAC TGT TGT CAG CAA CAT TAT AGC ACT C</td>
<td>2</td>
<td></td>
<td>1</td>
<td>GC TCA CGT TCG GGT CTG</td>
</tr>
<tr>
<td>V3 A4</td>
<td>ACT TAT TAC TGT CAG CAG TGG AGT AGT AAC CCA C</td>
<td>0</td>
<td></td>
<td>2</td>
<td>GC TCA CGT TCG GGT CTG</td>
</tr>
<tr>
<td>V3 A5</td>
<td>ACT TAT TAC TGT CAG CAG TGG AGT AGT AAC CCA C</td>
<td>3</td>
<td></td>
<td>3</td>
<td>GC TCA CGT TCG GGT CTG</td>
</tr>
<tr>
<td>V3 A6</td>
<td>ACT TAT TAC TGT CAG CAG TGG AGT AGT AAC CCA C</td>
<td>2</td>
<td></td>
<td>3</td>
<td>GC TCA CGT TCG GGT CTG</td>
</tr>
<tr>
<td>ΔH1;</td>
<td>GCC CTC AGG(T/G) TTC AGT GGC TGG GG(A/G) T</td>
<td>88/81</td>
<td>14/21</td>
<td>CAA/G CAA/G AAG TGG CTG</td>
<td>5</td>
</tr>
<tr>
<td>ΔP1</td>
<td>ΔV5 A7</td>
<td>GCC TGC AGG TTC AGT GGC GGT GGG T</td>
<td>88</td>
<td>14</td>
<td>CT GGG ACC AAC TGT GGT CTG</td>
</tr>
<tr>
<td>ΔV5 H4</td>
<td>GCC TGC AGG TTC AGT GGC GGT GGG T</td>
<td>90</td>
<td>t cg</td>
<td>16</td>
<td>G GGA GCA AGC TGG AGG TCA</td>
</tr>
<tr>
<td>ΔV5 H6</td>
<td>GCC TGC AGG TTC AGT GGC GGT GGG T</td>
<td>90</td>
<td></td>
<td>14</td>
<td>CTA GGA GCA AGC TGG AGG TCA</td>
</tr>
</tbody>
</table>

Average # of Deletions: All Sequences 16 ± 13 Non-aberrant Sequences Only 1.4 ± 2.5

Average # of Insertions: All Sequences 0.9 ± 2.3 Non-aberrant Sequences Only 0.9 ± 2.4

Average # of N Insertions: All Sequences 0.8 ± 2.2 Non-aberrant Sequences Only 1.0 ± 2.5

Average # of P Insertions: All Sequences 0.1 ± 0.6 Non-aberrant Sequences Only 0.1 ± 0.6
elements (79). Interestingly, 25% of these structures had nucleotide addition; four clones contain N insertions and one exhibited P addition. The incidence of N insertion is rare for the κ rearrangements of wild-type mice (3.2%) (78, 79), as TdT is generally no longer expressed in normal cells during recombination of this locus (77, 206). The presence of N addition in the SCID VJκ junctions suggests that κ rearrangement is occurring during the pro-B cell stage of development, perhaps simultaneously with rearrangement of the IgH locus. Alternatively, in the presence of defective DNA-PKcs, TdT activity may be improperly regulated either in terms of expression or function.

• Aberrant SCID κ Rearrangements Arise by Homology-Mediated Recombination

SCID clones corresponding to abnormally sized PCR products ΔVJκ4 and ΔVJκ5 show extensive deletions in both the Jκ and Vκ elements (Figure 4-6). The Jκ elements have between 14 and 21 nucleotides deleted while the Vκ elements have between 81 and 90 bp deleted. These deletions correspond to the expected size differences estimated from their PCR products. Three Vκ families contain a cryptic heptamer sequence [CAC(A/T)CT(C/T)] located approximately seven bases 3' of the region of identity in the aberrant VJκ junction sites (118). Recognition of this cryptic heptamer site may contribute to the large deletions observed in the Vκ elements of these structures.

The majority of the ΔVJκ4 and ΔVJκ5 sequences could have joined using homology, as five to seven bp at the joint can be found in both the Vκ and Jκ germline sequences. Interestingly, one ΔVJκ structures contains N addition (Clone H4). TdT is thought to interfere with the homology-mediated joining because rearrangements without N addition that occur in the presence of TdT rarely exhibit homology-directed joining (as reviewed in (28)). Clone H4 may simply represent the end product of a rearrangement event that used a cryptic heptamer sequence.
DISCUSSION

SCID MICE HAVE A RELATIVELY HIGH FREQUENCY OF NORMAL IgH & κ REARRANGEMENTS

SCID B220+ sIg- cells contain approximately 0.1-1% of the number of VDJH and VJκ rearrangements as wild type cells of that phenotype. These frequencies of SCID Ig rearrangement are likely an underestimation given the developmental differences existing between the C.B-17 and SCID B220+ sIg- cells (52% of the wild type cells are cμ-, compared with 100% of the SCID cells (206)). Previous studies determined that the frequency of coding joint resolution is 10^3-fold reduced in SCID cells (38), consequently the incidence of detectable SCID VDJH rearrangements is predicted to be a 10^6-fold lower than wild type. Given this prediction, the level of SCID IgH rearrangements detected here is unexpectedly high. Similarly, the ability to detect SCID κ rearrangements was also unanticipated given that SCID B cell differentiation is arrested at the pro-B cell stage of development where κ rearrangement rarely occurs in wild type B lineage cells (76, 78, 326). Indeed, initiation of κ rearrangement is undetectable by LM-PCR in non-IgH transgenic SCID mice (60, 327).

The level of VDJH and κ rearrangement in the B220+ and whole BM fractions do not reflect the degree of enrichment between the two populations (see Figures 4-1, -4, -5; Tables 4-2, -4). This finding may be explained by macrophage engulfment of presumably defective SCID B lineage cells (149). Analysis of the extent of IgH and IgL rearrangements in the B220- BM fractions was not assessed.

In general, individual SCID Ig rearrangements are indistinguishable from normal Ig junctions: With the exception of the aberrant Igκ rearrangements, the rearrangements have coding element deletions comparable to wild-type junctions. However, as a population, the SCID Ig joins contain an increased frequency of P addition (62% vs 2-15% in wild type (43)), a decrease in frequency of N addition,
and a greater incidence of homology mediated ligation than normal adult junctions. 86% of the IgH junctions analyzed and 44% of the non-aberrant κ joints are productively rearranged.

**SELECTION OF SCID B220+sIgM⁺ CELLS WITH VDJH REARRANGEMENTS**

B lineage cells derived from SCID mice lack sIgM, yet the VDJH structures isolated from them appear to have been selected: 86% were in frame, and maintained the invariant Cys residue at codon 92 that is critical for Ig folding (118). The ability to express functional Ig heavy chain in conjunction with surrogate light chain, has been shown to be critical in directing clonal expansion and differentiation from the pro-B to the pre-B cell phenotype (78, 171, 186). The lack of aberrant IgH rearrangement in SCID primary cells may reflect this critical need for a functional heavy chain, i.e. large deletions in the IgH rearrangements may prove deleterious for primary cells, thus accounting for the inability to detect such joins in these cells. (Only in the Igκ loci did we observe substantial deletions.) Moreover, positive selection of cells bearing resolved IgH coding joins provides a means of reconciling the relatively large degree of detectable SCID Ig rearrangements with the estimated 10³-fold reduction in SCID coding joint formation observed in recombination substrates (38).

Isolation of four different VDJH structures containing identical DJH junctions, and either different VH elements or different processing at the V-DJH border (Figure 4-3: clones A10, A16, A17, & A18), further suggests that SCID cells bearing a resolved coding join undergo cellular expansion. Generally, wild-type pro-B cells undergo one round of cell division post DJH rearrangement prior to V-DJH joining (95, 328). The occurrence of four different VH-D₂₉JH₂ junctions suggests that SCID B cell progenitors harboring this DJH rearrangement underwent at least two rounds of division and because of sampling probabilities, likely underwent many more
rounds of division prior to V-DJH joining. Since this DJH rearrangement is in RF1, one cannot attribute the apparent clonal expansion to ligand selection because only DJH junctions in RF2 can be expressed as a Dμ protein (83, 93, 185, 190). In addition, the number of unique SCID DJH rearrangements isolated is very limited compared to wild type, despite an estimated 4x10^5-7x10^5 SCID DJH rearrangements per femur (245). Collectively, these observations suggest that SCID progenitor cells bearing successful heavy chain rearrangements (i.e. DJH or VDJH), undergo several rounds of cell division. Thus, the limited repertoire and high occurrence of in frame IgH rearrangements in SCID B220+sIgM- cells suggest that these cells have undergone some form of selection, probably at a differentiation stage where the cells have substantial proliferative ability.

Given the lack of detectable κμ+ pre-B cells in SCID mice (J.P. and G. W. unpublished data (206, 207)), the mechanisms governing the selection of the productive SCID VDJH joins appear enigmatic. The inability to detect κμ+ cells is most likely a consequence of the frequency of SCID VDJH rearrangements and/or the level of κμ protein expression in these progenitor B cells occurring below the level of sensitivity of the FACS assays used. While the level of κμ protein expression may not be as high as that observed in wild type cells, a small number of heavy chain molecules may be sufficient to pair with the surrogate light chain to direct proliferation of these cells. Thus, the failure to detect κμ protein and the presence of an apparently "selected" IgH repertoire may not be irreconcilable observations.

If κμ+ cells are present in SCID mice, why then do allelically excluding IgL transgenes fail to rescue B cell development (172, 213, 327)? Several possibilities may account for this discrepancy: First, the IgL transgene may not be expressed early enough to prevent initiation of κ recombination in the SCID pro-B cell cells. Failure to resolve the resulting DSBs in the endogenous κ locus would culminate in
cell death. Secondly, SCID cells may attempt secondary κ rearrangements, despite the presence of a functional IgL transgene. Indeed, secondary κ rearrangements have been observed in wild type mice even after an in frame κ rearrangement has occurred (329-331). Similarly, SCID cells bearing productive IgH rearrangements may initiate secondary D to JH joining on their second allele, the inability to resolve these lesions would preclude B cell development in the presence of an IgL transgene. It is also possible that the IgL transgene may be unable to pair with the endogenous SCID IgH proteins.

Alternatively, other mechanisms independent of surface Ig expression may act to influence both the frequency and nature of detectable SCID Ig rearrangement. For example, 43% and 59% of the SCID VDJH and DJH joins, and 80% of the aberrant κ rearrangements showed evidence of joining by homology. Sequence identity between recombining elements likely serve to favour the resolution of the DSBs into a characteristic junction. Thus, a particular rearrangement may have arisen within independent progenitor cells. If so, however, one would expect to isolate the same Ig joins among different SCID mice - a phenomenon which is not observed.

Similarly, homology mediated resolution and/or other mechanistic biases inherent to the V(D)J recombinase may account for the limited number of deletions observed at the V-DH border of IgH rearrangements. Previously, the limited amount of deletion at the V-DH border in normal B cells was thought to reflect positive selection for cells that maintain the Cys92 residue that is essential for Ig folding (83). The DNA sequence encoding Cys92 is usually six base pairs from the 3' end of VH elements (332). However, constraints generated by DNA sequence homologies (82, 83) or by chromatin structure (144) may act to limit the number of deletions observed. This notion is supported by finding conservation of DNA encoding the Cys residue after recombination in extrachromosomal recombination
substrates that contain germline VH, DH and JH elements, but do not produce functional Ig protein (C. Yu, M. Larijani and G.W., unpublished data).

VH81X usage provides further precedence for Ig-independent selection of Ig repertoire: Although 20-40% of early B cell precursor VDJH rearrangements utilize VH81X, only 5% of mature peripheral B cells have this element on their surface (316, 333). B lineage cells derived from mice whose transmembrane region of the μ heavy chain has been disrupted (μMT) (176), also exhibit the same pattern of VH81X overusage and decline (316). Thus, although the mechanisms governing this skewing of the IgH repertoire remain enigmatic; these mechanisms appear to be unaffected by the scid defect.

How then do we account for the profound absence of B and T cells in SCID mice? The lack of sIg+ B cells in SCID mice could simply be a result of the IgH and Igκ rearrangements residing in different cells. The data presented here involve the analysis of bulk populations, not individual cells. As discussed above, introduction of fully functional single Ig gene rearrangements into SCID mice fails to rescue B cell development (170, 172, 213, 327). Thus, although light chain rearrangement can occur independently of IgH rearrangement and pBCR expression (78), the absence of increased leakiness in Ig transgenic SCID mice argues for an explanation other than the IgH and Igκ rearrangements residing in different cells.

Both wild type and SCID A-MuLV lines have the capacity to undergo secondary endogenous DJH and κ gene rearrangements (199, 237, 240, 315, 334, 335). It is tantalizing to speculate that such rearrangements may also occur in primary cells. Indeed, wild type mice transgenic for strong, allelically excluding IgH and/or IgL, also undergo endogenous DJH and κ rearrangements (unpublished data, (172, 327)). It is currently unknown whether such endogenous rearrangements are initiated prior to or concurrent with the Ig transgene expression in primary
cells. Nonetheless, these findings cause one to postulate that the accessibility of the RAG genes for the Ig loci may not be completely shut off after a functional Ig rearrangement is made. Similar attempted rearrangements in primary SCID progenitor B cells would ultimately prove deleterious, possibly accounting for the scid mutation's profound and continuing effect on lymphoid development even in the presence of Ig transgenes.

**IMPRECISE RSS CLEAVAGE IN SCID MICE**

Two unique VDJH joins were isolated from two different SCID mice that appear to have arisen from recombination events where cleavage occurred on the wrong side of the heptamer (see Figure 4-3, clones A24 and B7). This pattern may have resulted from the palindromic nature of the heptamer, with the RSS being perceived to be in the opposite orientation by the recombination machinery. This "misreading" of the RSS likely occurs by the same enigmatic process controlling VH gene replacement because there are no recognizable nonamer-like motifs within the JH elements (166, 168, 336).

The observed junctions may have arisen in two ways: First, as these structures lack detectable DH elements, they may have arisen by a direct VH to JH joining event (Figure 4-7A). In this scenario, in order not to violate the 12/23 RSS rule, the palindromic heptamer had to be read in its opposite orientation using a non-consensus nonamer 12 bp into the JH element. The resulting junction would be classified as a hybrid joint. Alternatively, these junctions may have arisen by an inversion event to the "reverse" RSS during D-JH rearrangement, generating a signal joint with the DH 12 RSS joined to the "reversed" 23 RSS. A subsequent deletional VH rearrangement to the DH 12 RSS in the signal joint would generate the observed structures (Figure 4-7B). This scenario requires SCID to make inversionsal joins, a process rarely seen in SCID (38), although clone B2 has an inverted DQ52 element,
FIGURE 4-7: Possible mechanisms by which SCID VDJH structures exhibiting cleavage on the wrong side of the heptamer can be generated. (A) Pseudo hybrid joint formation: The "reverse" RSS is represented as a striped triangle. The normal 23 bp RSS flanking the VH and JH elements are depicted by an open triangle. In order for a hybrid joint to occur without violating the 12/23 rule, the reverse RSS has to be perceived as a 12 bp RSS. Processing at the coding junction border is depicted as a squiggle. (B) Pseudo inversion and subsequent deletional rearrangement. An inversional rearrangement event between the 3' 12 bp RSS of a DH element and the reverse RSS (23 bp) generates a signal and coding joint. A deletional rearrangement of a VH element to the 12 bp RSS of the signal joint will result in the observed junctions. In the deletional rearrangement step, the reverse RSS would be treated like a coding end and would exhibit processing (Clones A24 and B7 Figure 4-3).
demonstrating such events can occur. The resulting structures from imprecise RSS cleavage are intriguing because the scid defect is not believed to affect the initiation of recombination i.e., alignment and cleavage at the RSS of the recombining elements (60, 134) Cutting on the wrong side of the heptamer has also been observed in single-stranded RSS substrates in the cell-free recombination system (126, 127). Both the in vitro recombination system and SCID B cells lack enzymatic components needed for complete and efficient V(D)J recombination. The presence of the normal components in wild-type mice may act to prevent such inappropriate cleavage events and/or their resolution.

Studies of rearrangements within the TCR δ locus have revealed other types of hybrid junctions in both wild-type and SCID mice that exhibit imprecise RSS cleavage and joining in violation of the 12/23 spacer rule (51). While the nature of the aberrant RSS cleavage in T cells is distinct from that presented in Figure 4-3, these findings suggest that the inappropriate cleavage at the RSS occurs independent of the scid defect.

V(D)J REARRANGEMENT IN CONTEXT OF THE MURINE SCID DEFECT

Collectively, rearrangements isolated from SCID mice exhibit a distinctive phenotype from wild type rearrangements, i.e. an overall reduced frequency in N insertion, and an increase in P addition and homology-mediated joining (Figures 4-3, -6, (45-47, 242-245)). Thus, it is unlikely that these rearrangements arose from a genetic reversion event. Insights into how these rearrangements arise is provided by analysis of other DNA-PK deficient mutants.

Ku80 deficient cells (Ku80−/−) are also defective in V(D)J recombination, however both coding and signal joint formation is compromised by the defect (146, 263, 268, 298, 299, 311). Moreover, the rare signal joins that do arise in these Ku80−/− cells exhibit deletions (146) compared to ~20% of SCID signal joins (38).
Based on these data, Ku was proposed to protect the DNA ends from nucleolytic degradation during the recombination process (8, 146, 267, 268, 277, 280). However, full length coding hairpin and signal ends are also found in Ku80-/- cells (298, 299, 311). These data suggest that the initial cleavage synaptic complex (CSC) in V(D)J recombination is a stable intermediate that requires both the Ku heterodimer and DNA-PKcs to make the coding and signal ends available for further processing (298, 311). Further supporting evidence of this model is derived from the inability of Ku70 depleted nuclear extracts to support in vitro signal joint formation (337).

Why then are the effects on V(D)J recombination so different between the two DNA-PK deficient mutants? Ku likely functions to modify the accessibility of the recombination intermediates within the CSC through its helicase activity and additionally by its recruitment of DNA-PKcs (298, 312). Subsequently DNA-PKcs may act to permit the docking and/or activation of enzymes necessary for the differential processing and resolution of the DNA ends (298). Thus, failure to alter the accessibility of DNA ends within the CSC, and ultimately to recruit DNA-PKcs would result in a profound deficiency in both signal and coding joint formation characteristic of Ku80-/- cells. Ku protein expression is not compromised by the scid defect (267, 277), and perhaps more importantly DNA-PKcs expression is still observed (301). Association of the mutant DNA-PK with the CSC may allow recruitment of the enzymes necessary for efficient signal joint formation, but not coding joints (298). Differential requirements for phosphorylation of the recruited enzymes by DNA-PK may also contribute to the disparity between coding and signal joint resolution in SCID mice.

The above model also provides a unifying explanation for the more severe deficiency in V(D)J recombination in the equine versus murine SCID defects (270). The equine scid defect also maps to the DNA-PKcs, however these cells are
defective in both signal and coding joint formation (270). The mutation responsible for the equine SCID defect has been identified as a truncation mutation that removes the entire DNA-PKcs kinase domain and results in the complete absence of DNA-PKcs (271). Thus, the difference in DNA-PKcs protein expression likely accounts for the more profound deficiency observed in equine SCIDs.

How do the coding junctions and imprecise signal junctions arise in SCID cells? It is possible that DNA ends escape the Ku/CSC complexes which fail to recruit the mutant SCID DNA-PKcs as a result of Ku's helicase activity. The association of the RAG proteins to the RSS may confer a small degree of protection to the escaped signal ends. Furthermore, signal ends have large regions of identity that may help align and stabilize the accessible ends to allow for their resolution, and may account for the unusual deletions observed in some SCID signal junctions (38). The low frequency of SCID coding joint resolution (<0.1% (38)) may also be accounted for by this model. The extent of sequence identity among coding ends is much less than among signal ends (118). Thus, alignment and stabilization of the coding ends is expected to occur at both a much lower extent and frequency than signal ends. In support of this model, 43% and 59% of unique SCID VDJH and DJH junctions respectively, and the vast majority of the aberrant κ joins involved joining by homology.

Many of the SCID IgH and Igκ junctions are indistinguishable from wild-type joins, involving N addition and no apparent homology mediated joining (see Figures 4-3,-6). Several scenarios may account for the rearrangements that do not appear to be join by homology. First, because only the final product is examined, regions of identity may have been used, but these were subsequently removed by nucleases; or perhaps intermittent or single base pairings are enough to stabilize the DNA ends. Alternatively, these rearrangements may have arisen by residual
catalytic activity of the murine SCID protein. As discussed above, the murine \textit{scid} defect is not a null mutation. Indeed the mutant DNA-PKcs protein is found at 1-10\% the level of wild type (8-10, 301), and can still associate with DNA (301). Alternatively, the enzymes recruited and/or activated by DNA-PKcs may still associate with the SCID protein and be operative, albeit a lower efficiency, thus resulting in the few normal junctions.
CHAPTER V

DISCUSSION
In this thesis I present data demonstrating that non-leaky SCID mice contain relatively high numbers of resolved Ig V(D)J joins. With the exception of the κ locus, aberrant sized rearrangements were not observed in SCID BM cells. Normal-sized SCID DJH rearrangements are found at 1-10% of the levels observed in wild type whole BM. Both normal-sized VDJH and κ light chain rearrangements are found at approximately 0.1-1% of wild type B220+sIg- cells. The incidence of these Ig rearrangements is likely higher given the differences in B cell lineage development existing between the C.B-17 and SCID populations analyzed (48% of the C.B-17 B220+sIg- cells are cμ+, while cμ+ cells are undetectable in SCID BM (206)).

The SCID Ig joins, with the exception of the aberrant κ rearrangements, are indistinguishable from wild type junctions with respect to the number of deletions observed. The majority of the unique SCID VDJH and DJH joins (86% and 55% respectively) are in frame, while 44% of the SCID non-aberrant κ joins are productively rearranged. As a population however, the SCID Ig junctions are distinguished from normal coding junctions by an increase in both the frequency and length of P additions. Moreover, these joins are more fetal-like in nature, exhibiting a decrease in N addition, and a greater incidence of homology-mediated joining when compared to junctions derived from normal adult mice.

The characteristic long P insertions, lack of N addition and increased homology-mediated joining in SCID recombination were first observed in Ig and TCR joins derived from leaky SCID mice (45, 242, 338). However, it was unclear to what extent these unique features reflected Ag selection and/or the rearrangement process in the presence of the scid defect. Specifically, Ag receptors encoded by rearrangements bearing large P insertions and/or homology-mediated joining may be autoreactive. In normal mice, cells bearing
such receptors would either be anergized or would undergo V gene replacement. Indeed, a κ coding joint bearing large P insertions was found on an excision product from a secondary κ rearrangement in a wild type mouse (331). Thus, the isolation of rearrangements from non-leaky SCID lymphoid cells bearing the same features as leaky SCID rearrangements strongly argues that these characteristics are a consequence of the recombination process in the presence of the mutant DNA-PKcs.

Thymocytes from non-leaky SCID mice also exhibit normal levels of nucleotide deletion, an increase in P insertion and a fetal-like repertoire in their TCR rearrangements (46, 243, 244, 322). This observation indicates that neither lymphoid lineage appears to express enzymes that can compensate for the deficiency in DNA-PKcs activity in SCID mice.

APPARENT DISCREPANCY WITH A-MuLV DATA

Early studies with SCID A-MuLV lines and LTBMCs showed that endogenous coding joint formation was greatly affected by the scid defect. Typically, these joins involved large deletions that removed most of the rearranging Ig gene segments (39, 47, 235-238). Moreover, recombination substrate transfections into SCID A-MuLV lines determined that the frequency of coding joint formation was reduced 10^3-fold compared to wild type cell lines (38, 241). Superficially, the data presented in Chapters III and IV seem to contradict these findings. However, these findings are not as incongruous as they first appear.

- Primary Cells Lack Large Deletions in Their V(D)J Rearrangements

The quality of the endogenous Ig loci rearrangements differ greatly between primary SCID lymphoid cells and transformed or cultured SCID lines. Rearrangement events isolated from SCID A-MuLV lines, LTBMC and
spontaneous T lymphomas characteristically involve large deletions of the recombining gene segments, often resulting in the deletion of the entire recombining segments themselves (39, 47, 235-238). Generally, the Ig and TCR rearrangements isolated from both leaky and non-leaky SCID mice appear normal in size (the aberrant SCID κ rearrangements are a notable exception) (data presented here, (45-47, 242-244, 322, 338)).

Why are the rearrangements in primary SCID lymphoid cells so different from LTBMC or transformed SCID cells? The absence of detectable aberrant IgH joins in primary SCID cells likely reflects the critical role IgH plays in directing B cell development. Large deletions in the IgH locus may prove deleterious to a cell during pro-B to pre-B cell transition, resulting in the elimination of the cell. Indeed, the vast majority of the IgH structures isolated from SCID BM cells are productive rearrangements that maintain the invariant Cys residue at position 92. Moreover, aberrant-sized VDJH rearrangements have only been observed in BM cells from SCID/(Eμ)bcl-2 transgenic mice (221). On the other hand, both normal and aberrant rearrangements are observed in the κ locus of SCID B220+sIg- cells, where only 44% of the non-aberrant κ joins isolated are in frame. These results suggest that the κ joins are not subjected to selection in the SCID pro-B cells. Taken together, these observations suggest conditions which promote cell survival transiently alleviate the necessity for productive Ig rearrangements, and permit the existence of aberrant-sized rearrangements. Thus, it is not surprising that only the κ locus exhibited aberrant rearrangements in primary B220+sIg- cells, since the products of IgL recombination are not subjected to selection during the pro-B cell stage of development.

Another contributing factor that may account for the disparity between rearrangements in SCID A-MuLV lines and primary cells lies in the capacity of A-MuLV transformants to maintain active V(D)J recombinase activity (315, 334,
SCID A-MuLV cell lines, like wild type transformants, can undergo secondary DH to JH, as well as V to Jκ joining (199, 237, 240). Thus, even if a rare "normal" coding joint were formed, continuous rearrangement of the same allele in context of the scid defect would eliminate the junction, contributing towards the grossly aberrant lesions observed in the SCID cell lines.

Wild type B cells bearing a μ transgene carry endogenous DJH rearrangements (172). It is unknown if these DJH joins occur prior to or concomitant with surface expression of the transgene. It is entertaining to speculate that attempted secondary DJH rearrangements within the IgH loci may contribute towards the lack of B cell development in both IgL transgenic and non-Ig transgenic SCID mice.

• Accounting for the Apparent Differences in Coding Joint Formation Frequency Between Primary and Transformed SCID Cells

The frequency of coding joint resolution in SCID mouse lymphoid cells has been estimated to occur at 0.1% the level of wild type cells (38). This estimate of coding joint resolution frequency predicts that DJH and VDJH joins occur 10³- and 10⁶-fold less frequently in SCID B cells than in wild type cells. Yet, the level of detectable DJH rearrangement occurs at 1-10% of the frequency of wild type whole BM (245, 246). VDJH and κ joints are found at approximately 1% of the level of wild type B220⁺sIg⁻ cells. These frequencies are considerably higher than those predicted by the recombination substrate data. The high level of Ig and TCR rearrangements in SCID mice probably reflects clonal expansion and/or selection of few cells with non-aberrant resolved coding ends. Indeed, the limited diversity characteristic of SCID Ig and TCR junctions is consistent with clonal expansion (45, 47, 242-246).

Given the lack of detectable μ protein (206, 207), the mechanism(s) governing the selection of the productive SCID IgH structures appear enigmatic.
The high level of productive VDJH joins cannot be selected by Ag selection, as they were derived from surface Ig- cells. The most probable explanation for the discrepancy between finding in frame VDJH rearrangements, but not \( \text{\textmu}^+ \) cells, lies in the sensitivities of the assays used to detect the IgH rearrangements and protein. The incidence of cells bearing productive VDJH rearrangements (~1%), is likely too low to be detected by surface and/or cytoplasmic FACS staining. Indeed, surface pBCR expression is difficult to observe even in pre-B cells from wild type mice (165).

The introduction of strong allelically excluding IgL transgenes into SCID mice fails to support complete B cell development (172, 213, 327). Several scenarios may account for this inability to rescue B cell development: First, the IgL transgenes may be unable to pair with the endogenous rearrangements. Alternatively, the IgL transgenes may not be expressed early enough to prevent initiation of \( \kappa \) rearrangement in the endogenous loci of the SCID pro-B cells. Consequently, these SCID cells would die due to a failure to resolve the resulting DNA lesions in the \( \kappa \) locus. Similarly, as discussed above, the SCID pro-B cells which contain a productive VDJH join may have unresolved DSBs in their second IgH alleles. Finally, the inability of transgenes to support B cell development may reflect unforeseen functions of DNA-PKcs that may occur downstream of successful DSB resolution. In this regard, it is important to remember that DNA-PKcs has the capacity to activate a number of transcription factors in vitro (reviewed in (234)).

Receptor-mediated selection of cells bearing DJH rearrangements can not account for the high frequency of DJH joins observed. Only rearrangements in RF 2 give rise to the D\( \mu \) protein (185). Surface D\( \mu \) expression however, is believed to result in the termination of IgH rearrangement and consequently in cell death (83, 93, 190). Indeed, counter selection against D\( \mu \) protein is also
present in the SCID DJH joins (Chapter III), as seen by the less than expected 33% frequency of RF 2 usage. "Bystander" expansion of DJH rearrangements on the second allele of cells bearing a productive VDJH join is also unlikely given that the incidence of SCID VDJH rearrangements is considerably less than DJH joins. Rather, the high level of DJH joining observed may result from abnormal clonal expansion of the few progenitor cells that contain a complete DJH junction that is independent of surface Ig. SCID BM contain an estimated 4.0-7.0 x 10^5 DJH rearrangements, yet only 1-10 unique DJH sequences are isolated from a single mouse (Chapter III). Further evidence supporting this theory is derived from analysis of the isolated VDJH structures in Chapter IV. Several different VDJH structures derived from a single SCID mouse contain the identical DJH border, but involved a different VH element and/or processing at the V/DH border. Thus, the high level of rearrangements observed suggests that these joins do not reflect the frequency of successful resolution, but rather clonal expansion of those cells bearing resolved junctions.

Lastly, precedence does exist for Ig-independent influences on the primary repertoire. A large proportion (20-40%) of progenitor B cells carry VH81X rearrangements, however only 5% of mature peripheral B cells contain this VH element in their surface Ig (316, 339). This decline in VH81X usage is also observed in B lineage cells derived from μMT mice, and thus occurs independent of surface Ig expression (316). The mechanisms governing this Ig-independent selection remains enigmatic.

RATIONALIZING REPERTOIRE RESTRICTION IN SCID MICE

SCID coding junctions share fetal-like repertoire qualities, i.e., they exhibit a general lack of N addition and an increase in homology mediated joining. Several factors can account for the limited repertoire observed in SCID mice. As
discussed above, cells containing a resolved coding junction may undergo cellular expansion, thus skewing the observed repertoire. Moreover, some cells appear to have undergone positive selection through the pBCR. This hypothesis is based on the finding 86% of the VDJH joins derived from B220+sIg- BM cells are productively rearranged, maintaining the critical Cys92 residue necessary for Ig folding. Interestingly, the non-aberrant κ rearrangements isolated from these SCID pro-B cells do not appear to have been subjected to selection as only 44% of these joins are productively rearranged.

Alternatively, certain architectural features of the coding DSB ends may promote their resolution in the context of the scid defect. It is likely that regions of homology help align the coding ends, and act to promote their resolution in the absence of functional DNA-PKcs. In support of this model, joining by homology is observed in 43% and 59% of all SCID VDJH and DJH structures respectively, and all the aberrant κ junctions.

DNA-PK, V(D)J RECOMBINATION AND B CELL DEVELOPMENT

As discussed in Chapter IV, the recruitment of DNA-PKcs to the CSC appears to be essential for efficient signal joint formation (298, 311). The features characteristic of murine SCID coding joins are likely a consequence of inefficient recruitment and/or activation of the enzymes involved in the resolution of the coding ends by the mutant DNA-PKcs. Indeed, this requirement for DNA-PKcs would explain the difference in severity between the murine and equine SCID V(D)J recombination defects, as equine SCID cells lack detectable DNA-PKcs (270, 271).

Overexpression of the bcl-2 transgene in murine SCID mice results in the accumulation of B lineage cells bearing a mature B cell phenotype, but lacking in sIg expression (221, 222). Despite the lack of sIg expression, intracellular IgH
protein has been detected in these cells (222). Moreover, similar B cell differentiation is observed in IgH transgenic RAG deficient/bcl-2 mice, but not RAG deficient/bcl-2 mice (222), suggesting that this differentiation is mediated by signals through the pBCR. Indeed, SCID/bcl-2 mice exhibit relatively high levels of both normal and aberrant sized IgH rearrangements, but very little κ rearrangement (221). Given that SCID/bcl2 B lineage cells express almost wild type levels of DNA-PKcs (309), it is possible that the increased level of SCID protein may influence the efficiency of coding end resolution. Alternatively, albeit not exclusively, the enhanced cell survival mediated by the bcl2 transgene in pro-B cells may provide enough time to permit the resolution of the unresolved DNA ends in the presence of the defective DNA-PKcs. Indeed, delay of DSB mediated apoptosis may provide an explanation for the increased survival of pro-B cells in SCID/p53 deficient mice (207, 227, 259). Assessment of the diversity among the IgH rearrangements found in these mice may prove to be insightful. However, such analysis should be accompanied by comparing recombination substrate analysis of a series of A-MuLV transformed lines derived from both SCID and SCID/bcl-2 mice. If the wild type levels of the mutant DNA-PKcs promote coding joint resolution, why do the SCID/bcl-2 cells lack slg? As proposed above, it is possible that initiation of V(D)J recombination is not properly regulated in the presence of the scid defect. Thus, unresolved DSBs in either the IgH or IgL loci or aberrant IgL joins would prevent slg expression in the SCID/bcl2 cells.

REMAINING ISSUES AND FUTURE DIRECTIONS

Chapters III and IV present data demonstrating that SCID BM cells contain DJH, VDJH and κ rearrangement. Thus, the absence of detectable slg positive cells in SCID mice despite the presence of these relatively high levels of Ig
rearrangement presents a conundrum. This circumstance may be a consequence of DNA-PKcs influencing events other than recombination. However, the apparent lack of IgM positive cells may simply be due to the VDJ and κ rearrangements occurring in different cells. The data presented in this thesis were performed on cell populations, and thus this issue could not be addressed. Use of single cell PCR analysis to examine the rearrangement status of both the IgH and IgL loci within the same cell would resolve this issue. Furthermore, analysis of cDNA from single cells would address whether or not the DH, VDJ and κ rearrangements are being transcribed as efficiently as wild type cells.

The data presented in Chapters III and IV raise questions with regard to the mechanism underlying the generation of the Ig rearrangements in the presence of the scid defect. Are these junctions a consequence of an insufficient and/or inefficient DNA-PKcs protein? Or, are alternative repair enzymes responsible for the SCID-like characteristics of these joins (i.e. the increased frequency of homology-mediated joining and P addition)? Is DNA-PKcs kinase activity absolutely required for V(D)J recombination process to occur (i.e. DSB resolution), or is the kinase activity required for downstream events, such as indicating to the cell that the DNA lesion has been repaired?

The above issues are not easy to assess, given that the precise role(s) of DNA-PK in V(D)J recombination and DSB repair are just now being elucidated. For example, DNA-PK mediated phosphorylation of repair enzymes could lead to their activation, or conversely, their inactivation and subsequent dissociation from the DNA-PK/DNA complex (273). DNA-PK kinase activity may also serve to inactivate proteins, for example transcription factors, whose activities might interfere with the repair process. It is worth noting that DNA-PK has been shown to phosphorylate and inactivate RNA polymerase I in vitro (313). Furthermore, DNA-PK appears to have a role in cell cycle control. DNA-PK can
phosphorylate p53 \textit{in vitro} \cite{11, 340}, however p53 is upregulated and presumed to be activated in SCID mice \cite{207}. Moreover, SCID cells appear to have normal cell cycle G1/early S and G2 checkpoints \cite{259, 261, 262}. Evidence demonstrating DNA-PK functions in a capacity other than DSB repair comes from analysis of DNA-PK deficient cell lines. Both SCID and Ku80-deficient cells subjected to irradiation arrest in G2, but exhibit impaired progression back into cell cycle \cite{262}. Thus, DNA-PK appears to elicit signals (either indirectly or directly) necessary for reentry into cell cycle \cite{262}. Unexpectedly, this G2 arrest activity appears to occur independently of DSB repair \cite{262}.

Furthermore, as discussed earlier, the murine \textit{scid} defect is not a null mutation. Although SCID mice contain mutant DNA-PKcs protein, it is expressed at 10-fold reduced levels as compared to wild type cells \cite{8, 9, 199, 301}. Consequently, it is difficult to access the precise contribution of the mutant DNA-PKcs to the SCID phenotype. A DNA-PKcs deficient mouse should help clarify the issues raised above. If DNA-PKcs is involved in processes other than the resolution of DSBs, one would expect to observe a more severe block in DSB repair, V(D)J recombination and lymphoid development than the SCID mutation. Indeed, equine SCID which lack detectable DNA-PKcs protein, exhibit a more severe defect in V(D)J joining than murine SCID \cite{270, 271}. Moreover, a DNA-PKcs null mouse may help to determine the mechanism underlying the normal recombination in thymocytes observed upon sub-lethal irradiation of SCID mice \cite{225, 226}.

DNA-PKcs kinase activity may signal events downstream of DSB resolution that play a critical role in lymphoid development. Thus, introduction of transgenes bearing various mutated forms DNA-PKcs back into DNA-PK null mice, will help to more fully understand the function of the DNA-PKcs. Moreover, this transgenic analysis may be useful in determining the molecules
and/or substrates with which the DNA-PKcs interacts. (Candidates include the XRCC4 gene, as well as p53 and various transcription factors).

Previously, I raised the possibility that unresolved DSB in the Ig loci may be responsible for the inability of single IgH or IgL transgenes to rescue B cell development in SCID mice. Secondary rearrangements of the \( \kappa \) locus has been documented in wild type mice (329-331). Several studies suggest that \( \kappa \) rearrangement is rarely initiated in non-IgH transgenic SCID cells (60, 172, 327), raising the possibility that rearrangement may be initiated on the second IgH allele, and that these unresolved DSBs are responsible for the inability of the IgL transgenes to rescue B cell development.

To address the above possibility, LM-PCR assays could be used to determine whether new and/or secondary endogenous IgH and \( \kappa \) rearrangements are initiated. This analysis should involve non-Ig transgenic, as well as IgH (M54) and IgL (Vk8) transgenic mice on both the SCID and C.B-17 backgrounds. Given the sensitivity of SCID cells to DSBs, it may be advisable to also utilize SCID/E(\( \mu \))\( bcl-2 \) transgenics to promote the survival of cells with unresolved DSBs. In order to obtain adequate numbers of cells for this assay, SCID and C.B-17 B220\(^+\)\( ckit^+\)sIg\(^-\) cells would be expanded in stromal cell cultures containing Il-7 for several days. Il-7 deprivation has been shown to induce B cell differentiation, and to permit V(D)J recombination (309, 341). The cells will be removed from Il-7, and then subjected to the following analysis: First, the endogenous Ig rearrangement status of the clone will be assessed. Secondly, genomic DNA will be isolated from the cells and subjected to both traditional LM-PCR assays (60, 132), and modified LM-PCR assays to look for secondary rearrangements on the same allele. The modified LM-PCR assays will involve locus specific primers embedded in the rearrangement events (i.e. a VH, DH or Vk specific primer), as well as the linker-specific primer. Sequencing of the
products derived from the modified LM-PCR will confirm that a secondary rearrangement event has been initiated in these cells. Finally, an aliquot of the cultured cells will be subjected to RT-PCR analysis to determine whether the rearrangements observed are being expressed. Primers that amplify the region of the DNA-PKcs gene believed to be responsible for the SCID defect can also be used during the RT-PCR analysis in order to determine whether a genetic reversion event has occurred.

In conclusion, the data presented in this thesis demonstrates that SCID mice are capable of undergoing successful V(D)J recombination. These data also suggest that the scid defect may have a profound and continuing effect on B cell development beyond a deficiency in V(D)J recombination. The approaches described in this chapter represent experimental means to further elucidate the role of DNA-PK in both V(D)J recombination and B cell development.
REFERENCES


182. Tsubata, T., and M. Reth. 1990. The products of the pre-B-cell-specific genes (λ5 and VpreB) and the immunoglobulin μ chain forms a complex that is transported onto the cell surface. *J. Exp. Med.* 172:973.


developmentally distinct T cell populations in severe combined immunodeficiency mice expressing a TCRβ transgene. *J. Immunol.* 150:1263.


autoantigen is the regulatory component of a template-associated protein kinase

dependent protein kinase is activated by nicks and larger single-stranded gaps. J.
Biol. Chem. 269:16684.

kinase is inactivated by autophosphorylation of the catalytic subunit. J. Biol.
Chem. 271:8936.

that are essential for cellular repair: a fundamental principle of apoptotic death.

Hendrickson. 1996. DNA-dependent protein kinase is a target for a CPP32-like

292. Song, Q., S. P. Lees-Miller, Z. Zhang, D. W. Chan, G. C. Smith, E. S.
Alnemri, G. Litwack, K. K. Khanna, and M. F. Lavin. 1996. DNA-dependent
protein kinase catalytic subunit: a target for an ICE-like protease in apoptosis.
EMBO J. 15:3238.

293. Song, A., S. R. Burrows, G. Smith, S. P. Lees-Miller, S. Kumar, D. W. Chan,
J. A. Trapani, E. Alnemri, G. Litwack, H. Lu, D. J. Moss, S. Jackson, and M. F.
Lavin. 1996. Interleukin-1 beta-converting enzyme-like protease cleaves DNA-

294. Rathmell, W. K., and G. Chu. 1994. Involvement of the Ku autoantigen in
91:7623.

DNA-binding protein antigen Ku recognized by autoantibodies from patients


closely resembling the human autoantigen Ku, recognizes single- and double-
strand transitions in DNA. J. Biol. Chem. 268:10546.


