Molecular cloning and characterization of hBRAG and other genes as potential regulators of RAG1 in the B cell lineage

by

Laurent Karl Verkoczy

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

© Copyright by Laurent Karl Verkoczy, 1999
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.
GENERAL ABSTRACT

Molecular cloning and characterization of hBRAG and other genes as potential regulators of RAG1 in the B lineage

Degree of Doctor of Philosophy, 1999.
Laurent Karl Verkoczy
Graduate Department of Immunology in the University of Toronto

The RAG genes are critical in lymphocyte development and in antigen receptor rearrangement. Finding other factors associated with RAG in expression will allow insight into the B cell developmental program and possibly into the regulation of lymphoid-specific components of the V(D)J recombination reaction. The goal of my study was to isolate and characterize novel genes by virtue of their co-expression with RAG1 and RAG2 mRNAs. Genes of this sort would be of considerable general interest because they may potentially fall into one of the following three non-mutually exclusive categories: 1) genes involved in the regulation of human RAG1 and RAG2, 2) unidentified components of the recombinase machinery or, 3) other novel pre-lymphocyte specific genes (based on their temporally associated pattern of expression with the RAGs). In this thesis, I develop a differential display-based screening approach to identify and characterize genes that are co-expressed with human RAG1 in the context of the B cell lineage. Using this approach, I have identified several candidate genes, both known and novel, that co-express with and possibly regulate RAG1 in a B cell-specific context. I show that one of these is a novel, B cell-associated type II transmembrane glycoprotein that may be important in B cell signaling. Based on these studies, it may be possible to identify several factors that are not only important factors in RAG regulation, but also in B cell development.
ACKNOWLEDGEMENTS

I am grateful to my supervisor, Dr. Neil Berinstein for his support and encouragement throughout the years. Although I chose the lab for fairly naïve reasons: a view of the CN tower, Blue Jays tickets (which I only ended up getting once anyway) and a tennis doubles partner in Niclas Stiernholm (we did win a couple of OTA tournaments together though), this choice turned out to be beneficial as it trained me to become an independent scientist. Dr. Berinstein showed a rarely found amount of patience with me and my experiments, especially in my first two years of “post undergraduate” immaturity (such as when I ran gels with stock concentrations or more alarmingly when I nearly burnt the lab down and.........well, you get the picture).

I would also like to thank the members of my supervisory committee, Dr. Susanna Lewis and Dr. Gillian Wu, for their sound criticism and advice. I cannot find the words to describe my appreciation to Gillian Wu, her friendship and her direct manner. I won't forget one particular instance at the beginning of my program. It was my first year in "the big city", and I did not have a computer to work on my first “recent advances in immunology” exam. Not only did she lend me a computer to work on, but when I came over to her place, she gave me a hat to wear on my way home, stating rather firmly that I was going to develop pneumonia the way I was dressed.

I thank my colleagues in the Berinstein lab who have all been supportive of my research, including the "RAG team", Ali Zarrin, Ivan Fong, and Jane Gillis, as well as others in the lab who have been extremely co-operative, including Barbara Guinn, Elena Iaknanina, and Abhijit Ghose. Thanks are in order to those who have helped near the end of my program, in particular the entire Reithmeier lab. Milka, Jeff, Jani, and Reinhart: thanks for all the advice, (technical and personal), and for including me as a regular in the Reithmeier coffee breaks. In retrospect, I think in the last year of my Ph.D. program, I spent more time in your lab than in the Berinstein lab! Recognition must also
be given to my parents and brothers in Regina for the continuous support they have afforded me over the years from a distance.

Most of all, I am indebted to my wonderful musician/scientist wife, Victoria Chiao. Vicki, not only did you largely help me with the editing and graphic aspects of this thesis, but you have had the patience, understanding, and faith to help me through the difficult times of my program. I am indebted to you. I am also indebted to my stepson, Alexander Chiao, and my two beautiful babies, Alyssa, and Laurent-Nicholas for bringing immense joy in my life and for making me laugh every day. To my family here in Toronto, I dedicate this thesis to you guys.

This work was supported in part by a Connaught scholarship, a Medical Research Studentship, and by NCIC grant # MT 10493. I would therefore like to thank these various funding agencies for providing me with the financial resources to do the numerous experiments presented in this thesis.
ABBREVIATIONS

aa           amino acids
Ab           antibody
Ag           antigen
AMuLV        Abelson murine leukemia virus
AP           arbitrary primer
APC          antigen presenting cell
ARE          A+U rich element
ATK          Abelson Tyrosine Kinase
bHLH         basic helix-loop-helix
BCS          bovine calf serum
BCR          B cell antigen receptor complex
BLAST        Basic Local Alignment Search Tool
bp           base pair
BSA          bovine serum albumin
BSAP         B cell-specific activator protein
Btk          Bruton's tyrosine kinase
C            immunoglobulin constant region
cAMP         cyclic-AMP
CD           cluster of differentiation
cDNA-RDA     representational difference analysis of cDNA
CDR          complementarity determining region
cDNA         complementary DNA
cpm          counts per minute
CR2          complement receptor 2
CRD          carbohydrate-lectin recognition domain
D            diversity gene segment
dBEST        Database of Expressed Sequence Tags
DD RT-PCR    Differential display RT-PCR
DEPC         Diethyl pyrocarbonate
DNA          deoxyribonucleic acid
DNA-PK       DNA-dependent protein kinase
DNase        deoxyribonuclease
dNTP         deoxynucleotide triphosphate
DN           CD4⁻CD8⁻ double negative
DP           CD4⁺CD8⁺ double positive
dsDNA        double-stranded DNA
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSBR</td>
<td>double strand break repair</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EBV</td>
<td>Epstein-Barr virus</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetra-acetic acid</td>
</tr>
<tr>
<td>EMSA</td>
<td>electrophoresis mobility shift assay</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>EST</td>
<td>expressed sequence tag</td>
</tr>
<tr>
<td>EtBr</td>
<td>ethidium bromide</td>
</tr>
<tr>
<td>FACS</td>
<td>fluorescence activated cell sorting</td>
</tr>
<tr>
<td>FCS</td>
<td>fetal calf serum</td>
</tr>
<tr>
<td>FDC</td>
<td>follicular dendritic cell</td>
</tr>
<tr>
<td>FISH</td>
<td>fluorescence in situ hybridization</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
</tr>
<tr>
<td>FR</td>
<td>framework region</td>
</tr>
<tr>
<td>GC</td>
<td>germinal centre</td>
</tr>
<tr>
<td>hBRAG</td>
<td>human B cell RAG-associated gene</td>
</tr>
<tr>
<td>HC</td>
<td>immunoglobulin heavy chain</td>
</tr>
<tr>
<td>Hck</td>
<td>hematopoietic cell kinase</td>
</tr>
<tr>
<td>HIM</td>
<td>hyper IgM syndrome</td>
</tr>
<tr>
<td>HMG</td>
<td>high mobility group</td>
</tr>
<tr>
<td>HRP</td>
<td>horseradish peroxidase</td>
</tr>
<tr>
<td>HSV</td>
<td>herpes simplex virus</td>
</tr>
<tr>
<td>la</td>
<td>invariant chain</td>
</tr>
<tr>
<td>Id</td>
<td>idiotypic</td>
</tr>
<tr>
<td>Ig</td>
<td>immunoglobulin</td>
</tr>
<tr>
<td>IL</td>
<td>interleukin</td>
</tr>
<tr>
<td>ITAM</td>
<td>immunoreceptor tyrosine activation motif</td>
</tr>
<tr>
<td>ITIM</td>
<td>immunoreceptor tyrosine inhibitory motif</td>
</tr>
<tr>
<td>J</td>
<td>joining gene segment</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kD</td>
<td>kilodalton</td>
</tr>
<tr>
<td>KIR</td>
<td>killer inhibitory receptor</td>
</tr>
<tr>
<td>KLH</td>
<td>keyhole limpet hematocyanin</td>
</tr>
<tr>
<td>κ</td>
<td>immunoglobulin kappa light chain</td>
</tr>
<tr>
<td>LC</td>
<td>immunoglobulin light chain</td>
</tr>
<tr>
<td>LPS</td>
<td>lipopolysacharride</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>λ</td>
<td>immunoglobulin lambda light chain</td>
</tr>
<tr>
<td>MAb</td>
<td>monoclonal antibody</td>
</tr>
<tr>
<td>MHC</td>
<td>major histocompatibility complex</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-N-morpholino-propane-sulfonic-acid</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
</tr>
<tr>
<td>MTN</td>
<td>multiple tissue northern</td>
</tr>
<tr>
<td>N</td>
<td>non-germline encoded nucleotides</td>
</tr>
<tr>
<td>NLS</td>
<td>nuclear localization signal</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>OS</td>
<td>Ommen Syndrome</td>
</tr>
<tr>
<td>PALS</td>
<td>perivascular lymphocytic sheaths</td>
</tr>
<tr>
<td>PBL</td>
<td>peripheral blood lymphocyte</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>pfu</td>
<td>plaque forming units</td>
</tr>
<tr>
<td>PKA</td>
<td>protein kinase A</td>
</tr>
<tr>
<td>PKC</td>
<td>protein kinase C</td>
</tr>
<tr>
<td>PMA</td>
<td>phorbol-12-myristate-13-acetate</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>PNGaseF</td>
<td>peptide-N-glycosidase F</td>
</tr>
<tr>
<td>PTK</td>
<td>protein tyrosine kinase</td>
</tr>
<tr>
<td>PVDF</td>
<td>polyvinylidifluoride</td>
</tr>
<tr>
<td>RAG</td>
<td>recombination activating gene</td>
</tr>
<tr>
<td>Rag</td>
<td>RAG protein</td>
</tr>
<tr>
<td>RAP-PCR</td>
<td>PCR by random amplification of polymorphisms</td>
</tr>
<tr>
<td>RLCS</td>
<td>restriction landmark cDNA scanning</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>RSS</td>
<td>recombination signal sequence</td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcriptase</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcriptase-polymerase chain reaction</td>
</tr>
<tr>
<td>SAGE</td>
<td>serial analysis of gene expression</td>
</tr>
<tr>
<td>SCID</td>
<td>severe combined immunodeficiency</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulphate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>slg</td>
<td>surface immunoglobulin</td>
</tr>
</tbody>
</table>
### Abbreviations cont'd

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSC</td>
<td>sodium citrate/saline buffer</td>
</tr>
<tr>
<td>ssDNA</td>
<td>single-stranded DNA</td>
</tr>
<tr>
<td>SSH</td>
<td>supression subtractive hybridization</td>
</tr>
<tr>
<td>SLC</td>
<td>surrogate light chain</td>
</tr>
<tr>
<td>SP</td>
<td>single positive</td>
</tr>
<tr>
<td>Taq</td>
<td>Thermus aquaticus</td>
</tr>
<tr>
<td>TCR</td>
<td>T cell receptor</td>
</tr>
<tr>
<td>TdT</td>
<td>terminal deoxynucleotidyl transferase</td>
</tr>
<tr>
<td>T&lt;sub&gt;H&lt;/sub&gt;</td>
<td>helper T cell</td>
</tr>
<tr>
<td>TIGR</td>
<td>The Institute for Genomic Research</td>
</tr>
<tr>
<td>TM</td>
<td>transmembrane</td>
</tr>
<tr>
<td>TNF</td>
<td>tumor necrosis factor</td>
</tr>
<tr>
<td>TSS</td>
<td>transcriptional start site</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
<tr>
<td>V</td>
<td>variable gene segment</td>
</tr>
<tr>
<td>XLA</td>
<td>X-linked agammaglobulinemia</td>
</tr>
</tbody>
</table>
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iii</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>v</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xvii</td>
</tr>
<tr>
<td><strong>CHAPTER 1: GENERAL INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>Overview of General Introduction</td>
<td>2</td>
</tr>
<tr>
<td>1.1 Expression and function of B lineage genes in mouse and man</td>
<td>3</td>
</tr>
<tr>
<td>1.1.1 Stem cells and development of the hematopoietic system</td>
<td>3</td>
</tr>
<tr>
<td>1.1.2 Human B cell ontogeny</td>
<td>5</td>
</tr>
<tr>
<td>1.1.3 General stages of B cell differentiation in mouse and man</td>
<td>5</td>
</tr>
<tr>
<td>1.1.3.1 Stages of B cell development based on Ig rearrangement and expression status</td>
<td>10</td>
</tr>
<tr>
<td>1.1.3.2 Stages of early B cell development based on cell surface differentiation markers and gene expression</td>
<td>11</td>
</tr>
<tr>
<td>1.1.4 General comparison of mouse and human B cell development</td>
<td>12</td>
</tr>
<tr>
<td>1.1.5 Pattern of RAG1 and RAG2 mRNA expression</td>
<td>14</td>
</tr>
<tr>
<td>1.1.5.1 General features of RAG1/2 mRNA expression pattern in and out of lymphocyte development</td>
<td>14</td>
</tr>
<tr>
<td>1.1.5.2 The two &quot;waves&quot; of RAG1 and RAG2 expression in antigen-independent phases of lymphocyte development</td>
<td>15</td>
</tr>
<tr>
<td>1.1.5.3 Expression of RAG mRNA in precursor cell lines in vitro</td>
<td>18</td>
</tr>
<tr>
<td>1.1.5.4 RAG mRNA expression in antigen receptor-expressing cells from primary lymphoid tissues and the third &quot;wave&quot; of RAG mRNAs</td>
<td>19</td>
</tr>
<tr>
<td>1.1.5.5 Re-expression of RAG1/2 mRNAs in peripheral lymphoid tissues: the fourth &quot;wave&quot; of RAG expression in GC subsets</td>
<td>19</td>
</tr>
<tr>
<td>1.1.5.5.1 GC formation and architecture</td>
<td>20</td>
</tr>
<tr>
<td>1.1.5.5.2 RAG expression in GC subsets</td>
<td>22</td>
</tr>
<tr>
<td>1.1.5.6 The RAG-deficient background in lymphocyte development</td>
<td>23</td>
</tr>
<tr>
<td>1.1.6 Classes of genes co-expressed with RAG mRNAs in antigen-independent and/or antigen-dependent phases of B cell development</td>
<td>26</td>
</tr>
<tr>
<td>1.1.6.1 Pre-BCR and BCR components: VpreB, λ5, Igα, Igβ</td>
<td>26</td>
</tr>
<tr>
<td>1.1.6.2 BCR co-receptors: CD19, CD20, CD21, CD22, and CD45</td>
<td>29</td>
</tr>
<tr>
<td>1.1.6.3 Stage-specific cell-surface differentiation markers</td>
<td>31</td>
</tr>
<tr>
<td>1.1.6.4 Co-expressed and/or associated genes involved in V(D)J recombination</td>
<td>32</td>
</tr>
<tr>
<td>1.1.6.4.1 Terminal deoxynucleotidyl transferase (TdT)</td>
<td>32</td>
</tr>
<tr>
<td>1.1.6.4.2 Potential, non-essential components</td>
<td>33</td>
</tr>
<tr>
<td>1.1.6.4.3 Rag-associated proteins</td>
<td>36</td>
</tr>
<tr>
<td>1.1.6.5 Co-expressed genes potentially involved in RAG regulation</td>
<td>42</td>
</tr>
</tbody>
</table>
1.2 Regulation of RAG expression: a cellular and molecular perspective

1.2.1 Cellular extrinsic and intrinsic influences regulating RAG expression
1.2.1.1 Regulation of RAG expression "waves 1 and 2" in antigen-independent phases of lymphocyte differentiation
1.2.1.2 Antigen receptor-mediated differential regulation of RAG expression "wave 3" in primary lymphoid populations
1.2.1.3 Signaling-mediated regulation of RAG expression "wave 4" in mature GC B cells
1.2.1.4 A role for RAGs in GC receptor selection/diversification rather than receptor editing
1.2.1.5 Factors which influence RAG expression throughout lymphocyte development

1.2.2 The three molecular levels of RAG regulation
1.2.2.1 Genomic organization and conservation of the RAG locus
1.2.2.2 Inter- and intra-species RAG2 message diversity
1.2.2.3 Level 1: Transcriptional regulation of the RAG genes
   1.2.2.3.1 Evidence for RAG transcriptional regulation
   1.2.2.3.2 Structural/functional characterization of RAG core promoters
   1.2.2.3.3 The ongoing search for tissue-specific transcriptional regulatory elements flanking the core RAG promoters
1.2.2.4 Post-transcriptional regulation of RAG mRNAs
1.2.2.5 Distinct pathways are involved in transcriptional/post-transcriptional control of RAG constitutive and inducible mRNA expression

1.2.3 RAG regulation is only one level regulating V(D)J recombination

1.3 Differential expression analysis: an approach for identifying novel B lineage developmental genes

1.3.1 General considerations for isolating differentially expressed genes
1.3.2 Classic (non PCR-based) methods for isolating differentially expressed genes
   1.3.2.1 Differential screening
   1.3.2.2 Microarray analysis
   1.3.2.3 Subtractive hybridization
1.3.3 PCR-based approaches for identifying differentially expressed genes
   1.3.3.1 Differential Display (DD RT-PCR) and the basic principle behind standard DD RT-PCR methodology
   1.3.3.2 Other comparative DD RT-PCR-like techniques
   1.3.3.3 Representational Difference Analysis of cDNA (cDNA-RDA)
   1.3.3.4 Serial Analysis of Gene Expression (SAGE)
1.3.4 Various ways DD RT-PCR or other differential expression techniques have been used to identify novel lymphocyte-specific genes and/or genes involved in lymphopoiesis
Table of Contents

1.4 Type II transmembrane proteins with characterized immune functions 95
   1.4.1 The TNF ligand family 95
   1.4.2 The LY49 family 97
   1.4.3 Miscellaneous other type II TM proteins with characterized immune functions 98

Thesis Summary 101

CHAPTER 2: Isolation of genes negatively or positively co-expressed with human Recombination Activating Gene 1 (RAG1) by Differential Display PCR (DD RT-PCR) 103

2.1 ABSTRACT 105
2.2 INTRODUCTION 106
2.3 MATERIALS AND METHODS 108
2.4 RESULTS 115
   2.4.1 Differential constitutive and inducible RAG expression in OCI LY8 variants 115
   2.4.2 Quantitation of differential RAG expression in OCI LY8 clonal cell variants and rationale for using OCI LY8-C3P and C3-A11N in DD RT-PCR 117
   2.4.3 Selection approach for choosing candidate RAG co-expressing genes 119
   2.4.4 Purification of single amplicons from heterogeneous amplicate mixtures 126
   2.4.5 Sequence analysis of purified amplicons 128
2.5 DISCUSSION 131

CHAPTER 3: Molecular characterization of hBRAG, a novel cDNA co-expressed with human RAG1 and encoding a type II transmembrane glycoprotein 137

3.1 ABSTRACT 139
3.2 INTRODUCTION 140
3.3 MATERIALS AND METHODS 142
3.4 RESULTS 150
   3.4.1 Selection of candidate cDNAs detecting RAG co-expressing transcripts 150
   3.4.2 cDNA cloning of hBRAG 150
   3.4.3 Analysis of hBRAG nucleotide and predicted amino acid sequence 152
   3.4.4 Expression pattern of hBRAG cDNA 156
   3.4.5 Southern analysis and chromosomal localization of the hBRAG gene 160
   3.4.6 A role for the hBRAG gene product in the regulation of RAG1 162
3.5 DISCUSSION 166
\textbf{CHAPTER 4: Biochemical characterization of hBRAG as a potential BCR signal-enhancing glycoprotein dimer that associates with phosphorylated proteins in resting B cells}  

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 ABSTRACT</td>
<td>172</td>
</tr>
<tr>
<td>4.2 INTRODUCTION</td>
<td>173</td>
</tr>
<tr>
<td>4.3 MATERIALS AND METHODS</td>
<td>175</td>
</tr>
<tr>
<td>4.4 RESULTS AND DISCUSSION</td>
<td>182</td>
</tr>
<tr>
<td>4.4.1 Characterization of hBRAG antibody specificity</td>
<td>182</td>
</tr>
<tr>
<td>4.4.2 Identification of hBRAG as a membrane-integrated, glycosylated protein</td>
<td>184</td>
</tr>
<tr>
<td>4.4.3 Analysis of hBRAG protein expression in lymphoid cells and tissues and identification of hBRAG disulfide-linked multimers</td>
<td>188</td>
</tr>
<tr>
<td>4.4.4 hBRAG intracellular and surface isoforms</td>
<td>190</td>
</tr>
<tr>
<td>4.4.5 hBRAG is associated with phosphorylated proteins in unstimulated cells that are dissociated and/or dephosphorylated in response to BCR signaling</td>
<td>194</td>
</tr>
<tr>
<td>4.4.6 α-hBRAG ligation alone generates modest increases in protein tyrosine phosphorylation and may also increase the degree of BCR-mediated phosphorylation upon co-ligation with anti-μ</td>
<td>199</td>
</tr>
</tbody>
</table>

\textbf{CHAPTER 5: GENERAL DISCUSSION}  

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Technical considerations and refinements in screening for RAG co-expressing cDNAs by DD RT-PCR</td>
<td>204</td>
</tr>
<tr>
<td>5.1.1 Pros and cons of DD RT-PCR relative to other techniques for assaying differential gene expression</td>
<td>204</td>
</tr>
<tr>
<td>5.1.2 An in-depth look at the limitations of DD RT-PCR and current proposed refinements/modifications to the technique</td>
<td>205</td>
</tr>
<tr>
<td>5.2 General structural and functional implications of hBRAG studies</td>
<td>210</td>
</tr>
<tr>
<td>5.2.1 The potential role of hBRAG in normal B cell development, RAG regulation, and V(D)J recombination: general comments</td>
<td>210</td>
</tr>
<tr>
<td>5.2.2 Structural comparison of hBRAG to other characterized type II proteins in the immune system</td>
<td>211</td>
</tr>
<tr>
<td>5.2.3 Speculations as to molecular mechanisms of hBRAG function in normal B cell development and RAG regulation</td>
<td>212</td>
</tr>
<tr>
<td>5.2.4 Proteins with homologies to hBRAG</td>
<td>215</td>
</tr>
<tr>
<td>5.2.5 hBRAG as a Rag-associated protein?</td>
<td>218</td>
</tr>
<tr>
<td>5.2.6 Future directions with respect to hBRAG</td>
<td>219</td>
</tr>
<tr>
<td>5.2.6.1 Further characterization of hBRAG protein and potential associations</td>
<td>219</td>
</tr>
<tr>
<td>5.2.6.2 Approaches towards understanding hBRAG function in RAG regulation and B cell development</td>
<td>220</td>
</tr>
</tbody>
</table>
5.3 Identification of BSAP as a RAG co-expressed candidate in DD screens

5.3.1 The potential role of BSAP and other upstream factors in B cell-specific RAG regulation

5.3.2 Emergence of an increasingly complex picture of RAG regulation: further evidence for distinct regulation of RAG expression in T and B lineages

5.3.3 An overall model of hBRAG, BSAP, and other potential regulatory factors involved in early and late B cell-specific RAG regulation

5.3.4 Implications for altered hBRAG and BSAP gene expression in potential B cell-specific immunodeficiencies, cancers, and autoimmunities

5.4 Concluding remarks

REFERENCES

APPENDIX
List of Figures

List of Figures

GENERAL INTRODUCTION (CHAPTER 1)

Figure 1.1 Hypothetical lineage relationship of hematopoietic cells 4
Figure 1.2 Overall model of B cell development 7
Figure 1.3 Expression profile of developmentally regulated genes in normal
B cell differentiation and developmental position of corresponding
human B lymphoid malignancies 8
Figure 1.4 Schematic representation of RAG expression
throughout lymphocyte development 16
Figure 1.5 Structure of a typical GC and model of RAG
expression and function in GC subsets 21
Figure 1.6 The RAG disruption in mouse lymphocyte development 24
Figure 1.7 Transcription factors in B cell development 28
Figure 1.8 Structure of the RAG proteins 38
Figure 1.9 Hypothetical model of signaling-mediated regulation of the four
RAG expression "waves" throughout B lymphocyte development 53
Figure 1.10 Three possible mechanisms of RAG-mediated secondary
rearangements for the purpose of "receptor editing" 57
Figure 1.11 Genomic organization of the human RAG locus 64
Figure 1.12 General outline of subtractive hybridization 79
Figure 1.13 The principle of differential display RT-PCR (DD RT-PCR) 81
Figure 1.14 The principle of Restriction Landmark cDNA Scanning (RLCS) 86
Figure 1.15 General schemes of the PCR-based subtraction methodologies known as cDNA
Representational Difference Analysis (cDNA-RDA) and Suppression Subtractive
Hybridization (SSH) 88
Figure 1.16 The principle of Serial Analysis of Gene Expression (SAGE) 89
Figure 1.17 Structural and functional aspects of type II TM proteins in immunology 96
LIST OF FIGURES cont'd

CHAPTER 2

Figure 2.1 Clonal relationship of OCI LY8 and variants 116

Figure 2.2 Quantitative comparison of RAG mRNA expression levels between OCI LY8-C3P parental cell line and clonally related variants 118

Figure 2.3 Differential display analysis to detect differentially expressed bands in the low RAG expressing clone OCI LY8-C3P and the high RAG-expressing variant C3-A11N 120

Figure 2.4 Reconfirmation of differentially expressed amplicates between OCI LY8-C3P and C3-A11N by Northern blot 123

Figure 2.5 Tertiary screens showing expression of transcripts detected by differentially expressed amplicates in unrelated RAG- and RAG' lines 125

Figure 2.6 Representative isolation (purification) of unique amplicons from heterogeneous DD RT-PCR amplicate mixtures 127

Figure 2.7 Summary of overall selection scheme for isolating candidate RAG co-expressing genes 135

CHAPTER 3

Figure 3.1 Combination of DD RT-PCR and northern blot identifies a 5.0 kb differentially expressed transcript in the high RAG-expressing variant C3-A11N 151

Figure 3.2 In vivo excised inserts from tertiary screens of Nalm-6 cDNA sublibraries 153

Figure 3.3 Map of hBRAG cDNA cloning strategy 154

Figure 3.4 Nucleotide sequence of the hBRAG cDNA and predicted amino acid sequence and secondary structure of the encoded protein 155

Figure 3.5 Co-expression of hBRAG and RAGl message in human early B cell lines 158

Figure 3.6 B-lymphocyte lineage specificity of hBRAG expression in human tissues 159

Figure 3.7 Genomic Southern blot analysis of hBRAG 161

Figure 3.8 Chromosomal localization of the hBRAG gene 163
LIST OF FIGURES cont’d

Figure 3.9 Increased levels of RAG1 mRNA in OCI LY8 cell lines transfected with the full length hBRAG cDNA 164

CHAPTER 4

Figure 4.1 Immunochemical characterization of rabbit polyclonal antiserum

Figure 4.2 N-Glycosylation of in vitro-translated hBRAG cDNA, and endogenous and transfected hBRAG products in cellular lysates 185

Figure 4.3 Western blot analysis of various human cell lines and tissues with α-hBRAG antibodies reveals hBRAG dimer formation 189

Figure 4.4 Analysis of hBRAG biosynthesis as assessed by immunoprecipitation of metabolically pulse and steady-state labeling of hBRAG A8-6P whole cell lysates 191

Figure 4.5 Analysis of hBRAG cell surface and intracellular expression 192

Figure 4.6 Detection of potential hBRAG-associated phosphorylated proteins in resting, but not BCR-ligated A8-6P lysates immunoprecipitated with α-hBRAG antibodies 195

Figure 4.7 Co-immunoprecipitation of A8-6P lysates with α-hBRAG, or candidate hBRAG-associated B-cell PTKs lyn and hck. 197

Figure 4.8 Analysis of tyrosine phosphorylation profiles in A8-6 whole cell lysates cross-linked with α-hBRAG, anti-μ, or both 200

CHAPTER 5

Figure 5.1 Possible hierarchy of factors involved in the early and late B cell-specific regulation of RAG transcription 228

APPENDIX

Figure A.1 Comparison of KIA0558 and hBRAG cDNA sequences 262

Figure A.2 Alignment of the KIAA0598 and hBRAG protein sequences 263

Figure A.3 Immunochemical characterization of rabbit polyclonal antiserum raised against N and C-terminal hBRAG peptides 264

Figure A.4 Reactivity of N and C-terminal hBRAG antibodies to in vitro-translated hBRAG cDNA products 265
LIST OF TABLES

INTRODUCTION (CHAPTER 1)

Table 1.1 Summary of rearrangement activity and gene expression through-out mouse B cell development 9

Table 1.2 Trans-acting components involved in V(D)J recombination: the eukaryotic recombinase machinery 35

Table 1.3 Extrinsic influences involved in inhibitory regulation (termination or downregulation) of RAG1 and/or RAG2 expression 49

Table 1.4 Extrinsic influences involved in positive regulation (induction or upregulation) of RAG1 and/or RAG2 expression 51

Table 1.5 Comparison of techniques for isolation of genes based on differential mRNA expression 76

Table 1.6 Summary of comparative techniques for identification of differentially expressed genes (RNA fingerprinting approaches) 84

Table 1.7 Examples of previously-unidentified genes with lymphocyte-specific expression patterns (and/or are potentially involved in lymphopoeisis) isolated by DD RT-PCR 90

CHAPTER 2

Table 2.1 Oligonucleotide primers used for DD RT-PCR 110

Table 2.2 Summary of DD RT-PCR (1ª screening) results for selecting genes co-expressed with human RAG1 121

Table 2.3 Summary of Northern blot analysis (2ª screening) for selecting differential display transcripts co-expressed with human RAG1 124

Table 2.4 Sequences of amplicons detecting human RAG1 co-expressing transcripts 129

CHAPTER 5

Table 5.1 Summary of limitations to analysis/selection of DD RT-PCR cDNA products and suggested refinements to original DD RT-PCR and subsequent screening steps 206
CHAPTER 1:

General Introduction
Overview of General Introduction

Based on the types of potential candidate genes that can be encountered in the overall screening strategy presented in this thesis, two major areas are covered in this introductory chapter: B cell development, and RAG expression/regulation. Furthermore, to serve as background for Chapter 5, in which the results of my actual screens and the molecular/biochemical characterization of hBRAG will be generally discussed, two other areas will also be briefly covered: 1) the various types of methodologies available for identifying differentially expressed genes, and 2) a brief review of the classes of type II transmembrane molecules with immune function. Although the RAG proteins are critical in V(D)J recombination, this process will only be discussed briefly with respect to RAG co-expressed and/or associated genes that may be involved in this process.
1.1 Expression and function of B lineage genes in mouse and man

A major focus of this thesis is concerned with identifying genes that co-express with RAG1 and RAG2 in B cell development. In this section, I will define B cell developmental stages and discuss the expression pattern of RAG1, RAG2, and other co-expressed genes in the context of the B cell lineage. For obvious reasons, normal mouse hematopoiesis and B cell development has been much more extensively characterized than in human both at the cellular and molecular levels. Most of the material reviewed below, however, will try to provide a general B cell developmental scheme unifying mouse and human development, but differences will be compared and pointed out throughout.

1.1.1. Stem cells and development of the hematopoietic system

Like other hematopoietic system members, B lymphocytes develop from multipotent stem cells in the adult bone marrow and fetal liver [1, 2]. The development of mature hematopoietic cells transpires from stem cells due to an assemblage of differentiation intermediates which undergo a progressive restriction in developmental potential (Figure 1.1; reviewed in [3, 4]). At one end of this spectrum are stem cells, which have extensive differentiative and self-renewal capacity and exist at very low frequencies in hematopoietic tissues (estimated at $1/10^4$-$10^5$ in bone marrow) [5]. At the other end are mature hematopoietic cells which have a limited life span, no self-renewal capacity, and require constant replenishment from the immature progenitor population. To study hematopoietic events, several assays which rely on the introduction of unique genetic markers, such as chromosomal aberrations or provirus integration into a specific stem cell have been devised to study hematopoietic ontogeny [1, 2, 6, 7]. The developmental potential of the clonable cell is then inferred from analysis of its clonal progeny after transplantation and expansion in a number of assay systems including irradiated hosts. The above assays have allowed the identification of various hematopoietic intermediates, one of these being a putative common lymphoid-restricted progenitor. Several observations submit corporeal evidence of a lymphoid restricted stem cell intermediate. The most substantial affidavit is the ability of some progenitors that have been present in long-term bone marrow cultures to reconstitute lymphoid cells in normal or SCID recipients in the absence of apparent myeloid reconstitution [8, 9]. Further studies using retroviral integration sites as unique genetic tags have not
Figure 1.1. Hypothetical lineage relationship of hematopoietic cells. The figure shows a model of hematopoietic development from multipotent stem cells (Sp). Sp, while maintaining the population size by self-renewal, are able to differentiate along the myeloid or the lymphoid pathway, leading to the generation of more restricted progenitors and eventually into mature hematopoietic cells. The myeloid lineages develop from a myeloid restricted stem cells (Sm) which give rise to CFU-s. CFU-s give rise to CFU-mix, which are the progenitors of bipotent and monopotent myeloid progenitors. The presence of lymphoid progenitors restricted stem cell (LC) has been suggested, but not proven. Mature T lymphocytes develop from pre-T cells in the thymus, which may also give rise to Natural Killer (NK) cells. B cells develop from pre-B cells in the adult bone marrow and fetal liver. More recently, a distinct bipotential fetal liver cell progenitor population (BM) has been described, which can give rise to B cells and macrophages in vitro (Cumano et al., 1992, reference [31]). Both Sm and LC have limited capacity for self-renewal relative to Sp.
yet given absolute proof of a lymphoid restricted stem cell [6, 7]. Another line of evidence suggesting a common progenitor is that mice with a targeted mutation in the Ikaros transcription factor show an absence of all lymphoid cell types while the myeloid lineages appear unaffected [10].

1.1.2 Human B cell ontogeny

Because the system I work with is human B cell lines and tissues, I will focus briefly on human B cell ontogeny. In general, pro-B and pre-B cells can be found in the fetal liver during week 8 of gestation, immature sIgM+ B lymphocytes are detected by week 9 of gestation, whereas mature sIgM+/sIgD+ B cells appear in week 12 (reviewed in [11]). Isotype-switched B cells soon follow, and plasma cell secreting IgG or IgA antibodies may be detected by 20-30 weeks of gestation [12]. During the second trimester, production of B lineage cells shifts to the bone marrow.

Most of the serum Ig in the fetus comes from the mother. Active transport of maternal IgG across the human placenta begins ~20-21 weeks of gestation, and the fetus acquires levels higher than that of the mother by birth [13]. This maternally derived IgG is catabolized with a $t_{1/2}$ of approximately one month so that the serum concentration declines after birth. It reaches lowest levels at 3-6 months, by which time the infant begins to make sufficient amounts of its own IgG, the levels of which slowly build up to adult values by ~6 years of age. Other IgG classes peak to adult levels at different ages. For example, serum IgA does not reach adult levels until puberty, whereas, secretory IgA and serum IgM adult levels are reached between one to two years of age [14].

1.1.3 General stages of B cell differentiation in mouse and man

Specific B cell developmental stages are defined in several ways by several mechanisms including cell size, growth properties, Ig gene rearrangement status, and the expression pattern of several B cell-restricted genes and cell surface markers (reviewed in [15]). In mouse, several nomenclatures of early B cell differentiation have been proposed (reviewed in [16]). Three of these are: 1) The Osmond system (based on cell size, cell cycle status, expression of the markers terminal deoxynucleotidyltransferase (TdT) and B220, and intracellular/surface heavy chain (HC) expression) 2) The Melchers/Rolink system, primarily based on HC and light chain (LC) rearrangement status, expression
of the pre-B cell receptor, and the markers c-kit and CD25 and 3) The Hardy system, based solely on multiparameter expression analysis of various B-cell specific surface markers, and whose chronological progression of fractions is suggested by the phenotypic changes of sorted cells cultured in vitro. In human, a scheme has been proposed by Cooper and colleagues. While all these nomenclatures are uniform in the division of immature and mature B cell fractions, there is variation in the criteria dividing early B cell fractions, in particular what constitutes progenitor B (pro-B) and precursor B (pre-B) cell subpopulations. New reagents, such as antibodies specific for components of the human surrogate light chain (SLC) have permitted recent attempts to unify various mouse and human B cell differentiation events. In this context, a scheme for human development analogous to the Rolink/Melchers mouse nomenclature has been recently proposed by Melchers and colleagues, in which many of the same markers (in combination with human-specific markers) and nomenclatures are applied to analogous human early B cell subpopulations (see Figures 1.2 and 1.3; reviewed in [17]). These authors have also compared the three mouse nomenclatures to each other. In Figure 1.2, I have attempted to represent a unified, comparative model of mouse/human B cell development incorporating all nomenclatures. In Figure 1.3 and Table 1.1, the protein and/or mRNA expression patterns of various B cell-specific markers found in mouse and human differentiation), is shown (reviewed in [11]). Throughout this thesis, B differentiation events will be discussed predominantly in relation to the Hardy and Melchers/Rolink nomenclatures.

Differences in cell surface marker expression and other differing features of mouse and human B development will be detailed below. It is important to note that the standardization of nomenclatures used across mouse as well as between mouse and human remains a controversial issue, since the individual schemes represent only partially overlapping developmental subsets. Additionally, such models individually, and together, present the inherent danger of suggesting a mandatory pathway for the development of mature B cells. It is not certain that a mature B cell must pass through each of these phenotypic stages. Further, a population of cells selected according to the expression of a certain marker is unlikely to be homogeneous for a particular stage of differentiation. Overall, however, these nomenclatures together present a valuable tool for following B cell development in vivo.
Figure 1.2. Overall model of B cell differentiation. A schematic representation of B cell development, indicating Ig gene rearrangement, cell size, cycling status, surface expression of pre-BCR/BCR proteins, and relative steady state population sizes is shown. GL=germline. VDJHp,n=cells containing productive and nonproductive heavy chain rearrangements, respectively. LCp, n=cells containing productive or non-productive light chain rearrangements, respectively. RF=reading frame. Crosses represent cell death. Curved arrows indicate cell cycling. Note that this is a unified model based on the word systems of Cooper, Osmond, and Melchers/Rolink (reference [11, 16]) and the letter system of Hardy (reference [22]), and as such may contain only approximately partially overlapping subsets. Population steady state sizes in a normal 4-6 week mouse according to Melchers/Rolink are shown at the bottom of the figure by square boxes. The approximate population numbers are as follows: pro-B cells, ~1 million; pre-BI cells, ~2 million; pre-BII cells, ~30-50 million; immature B cells, 20 million, and mature B cells, 500-1000 million. Note that the mature B cell pool is further subdivided into CD5+ and CD5- cells (shown), resting and activated, and plasma and memory (not shown).
Figure 1.3. Expression profile of developmentally regulated genes in normal B cell differentiation and developmental position of corresponding human B lymphoid malignancies. The top of the figure shows relative cell surface expression of various differentiation markers, as assessed by panning and/or flow cytometry. The cells representing developmental stages also have various intracellular and surface-expressed proteins represented schematically. The bottom of the figure shows mRNA expression as assessed by RT-PCR. The nomenclature is adapted from a combination of Cooper and Melchers/Rolink (references [11, 17]). The Melchers/Rolink nomenclature is the same for human and mouse but is based on slightly different markers; see Figure 1.2 for comparison. Solid lines indicate expression patterns and the levels of expression are indicated by the thickness of lines. A legend is also included to the left denoting the symbols for the various pre-B and BCR components (which are shown on the cells cytoplasmically and/or surface expressed) as well as for the acronyms of various B lymphoid malignancies occurring through different stages of B cell differentiation.
## Table 1.1
Summary of rearrangement activity and gene expression throughout mouse B cell development

<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Developmental Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>early pre-B</td>
</tr>
<tr>
<td></td>
<td>early pre-B</td>
</tr>
<tr>
<td>Mismatch/Rejoin</td>
<td>A4 (pro-B)</td>
</tr>
<tr>
<td>IgH genes</td>
<td>IgH-D</td>
</tr>
<tr>
<td>IgL genes</td>
<td>VHD</td>
</tr>
<tr>
<td>mRNA expression</td>
<td>RAG1/2</td>
</tr>
<tr>
<td>Receptor components</td>
<td>VpreB and L5</td>
</tr>
<tr>
<td>mb-1 (lg-c)</td>
<td>-</td>
</tr>
<tr>
<td>5-29 (lg-d)</td>
<td>-</td>
</tr>
<tr>
<td>5q germinal transcript</td>
<td>-</td>
</tr>
<tr>
<td>y chain</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>-</td>
</tr>
<tr>
<td>Surface protein marker expression</td>
<td>B220 (CD45)</td>
</tr>
<tr>
<td></td>
<td>CD19</td>
</tr>
<tr>
<td></td>
<td>CD23</td>
</tr>
<tr>
<td></td>
<td>CD1</td>
</tr>
<tr>
<td></td>
<td>CD19 (II-29G)</td>
</tr>
<tr>
<td></td>
<td>CD79a</td>
</tr>
<tr>
<td></td>
<td>CD79b</td>
</tr>
<tr>
<td>Surface protein complexes</td>
<td>&quot;early&quot; SL complex (S7.4-S5-preB30-II/III)</td>
</tr>
<tr>
<td></td>
<td>pre-B rec. SL complex (p-VpreB/L5)</td>
</tr>
<tr>
<td>Intracellular protein expression</td>
<td>CD79b</td>
</tr>
</tbody>
</table>

1. Actual gene expression profiles have been done according to Hardy's nomenclature and all other schemes are compared relative to this. Note that comparisons between subfamilies of each scheme are only approximations, in reality there is likely overlap. The number of pluses represents the relative level of expression (within the same gene but not between genes). This information is adopted from a number of sources including [1, 2, 3, 4, 5, 6](references [16, 23, 42]).

2. Plasma cells in this scheme are included as a subset of mature B cells for eg: Bim or Bcl-2 is only expressed in activated, differentiated plasma cells (Henderson and Colman, 1988; reference [18]).

3. NA = not assessed.

4. + indicates occasional or trace activity or protein expression.

5. - indicates that expression of the gene is assumed to be negative based on overall expression pattern, even though it has not been assessed.

6. Gene expression is determined by RT-PCR.

7. + indicates possible expression in this compartment under proper induction events in GC or mantle.

8. Expression patterns for B cell transcription factors (except for B12, B67, B1, and BSAP) are only approximate and relative levels have not been determined (Hagman and Grosschedl, 1994; reference [43]).

9. Surface molecular expression as determined by flow cytometry.
1.1.3.1 Stages of B cell development based on Ig rearrangement and expression status

Crucial for B lymphopoeisis is proper expression of the Ig HC and LC genes, which are at least partly regulated by a plethora of tissue specific cis-acting promoters, enhancers, and transcription factors (reviewed in [18]). Ig gene transcriptional regulatory elements include the B gene associated proximal promoters, the Ig HC gene intronic and 3' enhancers (Eμ and 3'EH, respectively), the κ intronic and 3' enhancers (3'Ek and iEk, respectively), and the λ LC enhancer (EL). Tissue and developmental specificity of the Ig gene expression is not only governed transcriptionally, but also by the general order of Ig assembly during B cell differentiation. Since distinct temporally-regulated Ig rearrangement and expression events occur through B lymphopoeisis, B cells can be divided into four general differentiation stages based on these events: progenitor (pro-B), precursor (pre-B), immature B, and mature/plasma cells.

The process mediating Ig gene rearrangements, V(D)J recombination, is initiated during the pro-B cell stage by a D-JH recombination event at the HC locus. Pro-B cells do not express Ig proteins but may actively transcribe their Ig HC gene segments before rearrangement. These sterile transcripts may either contain isolated VH gene segments or Cμ domains or can initiate upstream of the JH proximal DHQ52 gene segment, transcribe through JH, and splice appropriately to Cμ. Germline transcription is thought to reflect the “accessibility” of an Ig locus to the recombination enzymes involved in V(D)J joining (discussed further later on). Promoter sequences are found upstream of most DH gene segments, and DJCμ transcripts are easily detected. In the Osmond classification system, pro-B and pre-B cells are divided by the absence and presence of cytoplasmic μ expression (the product of successful rearrangement of a VH segment to a previously joined DJμ) [16]. Thus, pre-B cells by this classification are earliest cells of the B lineage to produce conventional HC protein, but are devoid of cell surface immunoglobulin (sIg) expression since LCs are not synthesized until Ig rearrangement processes in the κ of λ loci are completed. During the final stages of primary B cell development, the pre-B cell rearranges the V and J elements at one of its LC loci. Following successful V to Jκ or V to Jλ gene rearrangement, IgM monomers are expressed on the surface of the immature B cell, non-covalently associated with Igκ and Igβ. This multimeric complex serves as the so called B cell
receptor (BCR) for antigen, with specificity provided by the Igs, and signal transduction capabilities provided by Igα/β (reviewed in [19]). Immature B cells are sensitive to clonal deletion if their receptor recognizes multimeric self-antigens with sufficient avidity. The next modification of Ig expression by B lineage cells is the production of cell surface IgD (sIgD), endowing the mature B cell with two classes of antigen receptors, both with the same specificity. The mature B cell can be triggered to proliferate and differentiate in either a T cell-independent fashion, by antigens such as carbohydrates that contain multiple repetitive antigenic determinants, or with help of T cells and their cytokine products (reviewed in [20]).

1.1.3.2 Stages of early B cell development based on cell surface differentiation markers and gene expression

In the context of cell surface differentiation markers, pro-B cells can generally be defined as the first cells committed to B lymphopoeisis based on their expression of CD45RO (B220), CD19, and CD43. Coinciding with HC rearrangements, a decrease in CD43, lack of TdT, and the appearance of the pre-B cell receptor complex (pre-BCR) can all be used as criteria to distinguish the initiation of the pre-B cell stage. Early large cycling pre-B cells are identified by decreased RAG1/2 expression. As these pre-B cells continue to differentiate, the RAGs are induced and LC rearrangements begin. Antigen-independent B cell development concludes by expression of sIgM (signifying successful LC rearrangements). The mature cells emigrate from the bone marrow into the periphery and can be detected by increases in B cell-specific markers such as sIgD, CD21 and CD22.

The various aforementioned individual classification schemes all have unique ways of further dividing populations of pro-, pre-, or mature B cells using various B-specific and/or stage-specific differentiation markers. In particular, the pro-B cells of Osmond’s scheme are subdivided into three populations based on TdT expression and the Osmond cytoplasmic μ+ pre-B cells are further divided into large pre-B cells and pre-B cells according to their size [16]. With respect to Melchers/Rolink’s original mouse nomenclature [21], two other surface markers are used to subdivide pre-B cells, c-kit (the stem cell factor receptor), and CD25 (the IL-2Rα chain): pre-BI cells are defined as ckit CD25 (in addition to being TdT, and cytoplasmic μ+), and pre-BII cells are c-kit CD25 (in addition to being TdT
and cytoplasmic $\mu^+$). Likewise, Hardy has proposed a subfractionation scheme for pro- and pre-B cells (fractions A-D) based on the cell surface markers B220, CD43, BP-1, and heat stable antigen (HSA) [22]. More recently, fraction A has been even further subdivided into three subpopulations, $A_0$, $A_1$, and $A_2$, where $A_0$ cells represent putative lineage-unrestricted pro-B cells [23]. Analogous to mouse early B cell differentiation, further subfractionation is achieved in human germinal centre (GC) mature B cells with the surface markers CD23, CD38, and CD77; this will be detailed further in the context of RAG regulation at this stage [24].

1.1.4 General comparison of mouse and human B cell development

Overall, B cell development in mouse and human development reveal striking similarities, including the general temporal sequence of B cell developmental events including Ig rearrangements and B cell-specific and/or stage-specific gene expression patterns. Furthermore, most B-cell specific and/or stage-specific genes are well-conserved between the two species and their mutant phenotypes generally reveal parallel effects in B cell development. This is best exemplified by two recent observations. First, patients bearing mutations in the $\lambda5/14.1$ gene, have an agammaglobulinemia and B cell deficiency that is strikingly similar to $\lambda5^+$ mice [25]. Similarly, hyper-IgM syndrome (HIM) patients, who have mutations in CD40L (resulting in either reduced CD40L expression or a version of CD40L incapable of binding to CD40 on B cells), conspicuously resemble CD40L$^+$ mice [26]. The developmental block in both occurs in peripheral B cell pools, which are rendered largely non-functional due to defective GC formation, isotype switching, affinity maturation, and plasma and memory cell differentiation [27-29]. As a result, HIM patients and CD40L$^+$ mice have a severely compromised humoral immune system, in particular, high levels of IgM but markedly reduced levels of other Ig isotypes.

Despite these overall similarities, there are some notable differences. One obvious difference is homeostatic considerations, including the markedly larger size of B cell compartments and slower ontogeny kinetics in human (reviewed in [30]). Another is the differential effect of some cytokines on mouse and human B cell development. For example, IL-7, a 25 kDa soluble growth (and possibly differentiation) factor, can support proliferation of mouse pro-B in the presence of stromal cells in vitro.
(reviewed in [4]), but it has not been possible to grow the corresponding pro-B cell human subsets with IL-7 in vitro, either in the presence of stromal cells, or together with any other cytokines. IL-7 also has differential effects on target genes it regulates in man and mouse: in particular acts in a downregulatory manner on certain genes such as RAG and TdT in human pre-B cells, whereas in mouse, IL-7 upregulates or promotes sustained expression or upregulation of these genes [31]. Furthermore, IL-7 seems to be of lesser general importance in human than mouse B cell development, since patients with an inactive IL-7 receptor γ chain (IL-2Rγ) appear to have normal B cell development [32], while IL-2Rγ− mice have a complete block in the Melchers/Rolink pre-B1 to pre-BII transition [33]. Another difference involves mutations in the Btk (Bruton's tyrosine kinase) gene resulting in markedly different phenotypes [17]. The human mutation causes X-linked Agammaglobulinemia (XLA), an immunodeficiency in which there is a nearly complete block at the pre-B cell stage [34, 35]. In mouse the effect is XID, a far milder immunodeficiency in which there is only a partial developmental block in pre-B cells resulting in reduced but detectable mature B cells and serum Ig, suggesting the presence of a mouse-specific compensatory pathway/factor [36]. Finally, several markers thought critical for mouse B cell ontogeny are not expressed on human pro- and pre-B cells including: CD2, CD45RO (B220), c-kit, CD25 (IL-2Rα chain), CD43, and BP-1 [17]. Conversely, significant markers for B-cell development in man, such as CD10 and CD34, are absent in the mouse.

More subtle developmental differences also exist. Whereas diversification by TdT-mediated N additions appears restricted to the HC rearranged locus in mouse; in humans it occurs at both HC and LC [11]. Second, a more even ratio of κ to λ distribution is seen in human than in mouse where the κ isotype predominates. Third, a conserved ATG translation start site exists upstream of the second of three potential Dμ reading frames (rf2) and a truncated Dμ protein has been described in virally transformed mouse pre-B cell lines [17]. Open reading frames are uncommon upstream of human DJCμ transcripts, and Dμ proteins have not been described in humans.
1.1.5 Pattern of RAG1 and RAG2 mRNA expression

A major aim of this thesis is to further the molecular understanding of differentiation events occurring when the recombinase machinery is on or off in B cell development by identifying and studying the co-expression of genes, novel or known, co-ordinately expressed and/or regulated with RAG mRNAs. Therefore, to correctly understand the overlapping and/or co-ordinate expression of other genes in relation to RAG mRNA expression in the B cell lineage, the complex developmental mRNA expression pattern of the RAGs throughout B cell development will be discussed below.

1.1.5.1 General features of RAG1/2 mRNA expression pattern in and out of lymphocyte development

In general, RAG1 and RAG2 mRNA expression is tightly controlled: the genes are expressed in a concordant, lymphocyte- and temporally restricted fashion [37]. Interestingly, both in vivo and in vitro, RAG1 mRNAs are generally expressed 10-100 fold more abundantly than RAG2 mRNAs. Furthermore, extrachromosomal substrate assays performed in both lymphocyte and RAG-transfected non-lymphoid cell lines (mouse and human) have invariably demonstrated that RAG1/2 expression strongly correlates with levels of recombinase activity [38]. In vivo, the highest levels RAG mRNAs are expressed in primary lymphoid organs: fetal liver, bone marrow and thymic cortex [39]. Further subfractionation of cellular differentiation stages within the B and T cell lineage has demonstrated that RAG1 and RAG2 mRNA expression is first observed in mouse committed B and T progenitor lymphocytes undergoing their first series of V(D)J rearrangements: CD44+IL-2R- DN (CD4-CD8+) thymocytes and Hardy fraction A pro-B cells [40-42].

Initially, RAG expression was thought to be continuously expressed and only in pro- and pre-B and T differentiation stages [43, 44]; i.e. lymphocyte populations undergoing active rearrangements at thymocyte β and α loci (DN and DP TCR - thymocyte populations) and at B cell HC and LC loci (i.e. pro- and pre-B B220+IgM Hardy fractions B-D) after which developmental downregulation of RAGs then occurs. Recent studies, however, have suggested that the developmental window in which RAG mRNA expression occurs in the lymphocyte lineage may be more complicated than originally thought. These more recent findings will be discussed in detail below. A scheme for the relative amounts of
RAG mRNA expression throughout mouse lymphocyte development is shown in Figure 1.4.

While RAG1 and RAG2 mRNAs are usually co-expressed throughout lymphocyte development, some notable exceptions to the concordant and lymphocyte-restricted pattern of expression exist. Firstly, low but detectable levels of RAG1 expression alone have been detected by PCR in post-mitotic neurons of the mouse CNS [45], and in unfractionated mouse testes and ovaries [46]. Since sensitivity limits of Northern blotting (or RT-PCR) in such situations may be an issue, it cannot be ascertained if they represent genuine discordant expression of RAG1, or alternatively if RAG2 is co-expressed with RAG1 at lower relative levels, as is seen in other tissues. Significant expression of RAG1 alone has also been reported in a transitional SP thymocyte population [47] and high levels of RAG2 expression only have been reported in avian B bursal cells [48]. In human tissues, trace but detectable levels of RAG1 have also been detected in several non-lymphoid tissues: adult cerebellum, hippocampus, testis, kidney, and adult/fetal liver (see Chapter 3, Figure 3.6 and [49]; RAG2 was not assessed in these studies). While it is not clear what the roles of solitary RAG1 or RAG2 expression may have (if any at all), it has been speculated that RAG1 may be involved in a diversity-generating, rearrangement-like mechanism similar to V(D)J lymphocyte recombination in the neuronal lineage. Since RAG1<sup>−</sup> mice do not have any observable CNS defects [50], it remains to be determined if the RAG locus has a role in neuronal development and/or diversification. RAG2 alone has been implicated in the Ig-specific gene conversion process documented in B cells in the chicken bursa of Fabrecius [48]. However, its requirement in this process is questionable based on the analysis of the avian bursal-derived DT40 cell line which has its RAG2 gene disrupted but retains the ability to undergo gene conversion [51].

1.1.5.2 The two "waves" of RAG1 and RAG2 expression in antigen-independent phases of lymphocyte development

In B cell differentiation, two "waves" of RAG expression have been demonstrated whose peaks coincide with HC and LC rearrangements (Figure 1.4A; [23, 52-54]). Specifically, applying Melchers and Osmond scheme of B cell development, the first wave of RAG expression is induced in early pro-B cells (Figure 1.4A). RAG expression levels remain high throughout early and late pro-B cell stages,
Figure 1.4. Schematic representation of RAG expression throughout lymphocyte development.

For each developmental stage, *in vivo* and *in vitro* regulators of RAG expression are shown. Question marks denote potential regulators for which there is currently no direct evidence. The B cell developmental scheme is represented according to the nomenclature of Hardy (reference [22]) and the T cell developmental scheme is represented according to Godfrey and Zlotnik, 1993. The αβ T cell lineage only is shown, but γ and δ chain rearrangements both occur roughly around the β chain rearrangement period (reference [41]). Dashed lines represent possible expression or rearrangement. Note the coincidence of RAG expression and rearrangement activity. The relative levels of RAG expression shown is arbitrary and is loosely based on the studies of Melchers, Hardy, and McDonald for the first two waves of RAG expression (references [41, 42, 53]) and those of Nemazee and Nussenzweig for the third and fourth waves of RAG expression (references [82, 228, 252]). The developmental expression pattern of RAG expression in early T cell development is analogous to that in early B cell development; however, less is known about RAG expression in immature and mature T cell development. In particular, it is currently not clear if a third wave exists in T cells or if expression persists until TCR crosslinking shuts off RAG mRNA. A fourth wave of RAG reexpression in CD4⁺ T cells has been described by McMahon and Fink, 1998 (reference [74]), analogous to re-expression in GC B cells.
Wave 1
- IL-7
- Ppi, Pp2 phosphatase activation
- BCR signaling
- Pre BCR signaling
- Pax-5?
- E2A?

Wave 2
- IL-7
- Ppi, Pp2 phosphatase activation
- BCR signaling (high avidity)
- Pax-5?
- E2A?

Wave 3
- Ppi, Pp2 phosphatase inhibition
- BCR signaling (low avidity)
- Pax-5?
- E2A?

Wave 4
- Ppi, Pp2 inhibition
- BCR signaling
- BMAG

Relative RAG1/2 mRNA expression/
Potential regulatory influences

Developmental stage (Hardy)

Rearrangement status
Ig heavy chain
Ig light chains

A

B

Relative RAG1/2 mRNA expression/
Potential regulatory influences

Developmental stage

Rearrangement status
TCR β chain
TCR α chain
in which HC D-JH and V-DJH rearrangements occur, respectively. Subsequently, they differentiate into large pre-B cells that express pre-B receptors: i.e. \( \mu \) chains in association with the SLC (consisting of \( \lambda5 \) and VpreB proteins, which will be detailed later) and at this stage, RAG expression is sharply downregulated [53, 54]. It has been proposed that downregulation of RAG1/2 expression after the completion of a productive rearrangement at one HC allele might guarantee allelic exclusion i.e. prevention of further rearrangement on the other allele. This signal that may come from the pre-B cell receptor itself [15]. Nevertheless, in the subsequent small pre-B cell stage, the pre-B receptors disappear and the second "wave" of RAG expression is induced, which participates in V-J recombination of LC genes. In the scheme proposed by Hardy, the sharp decrease in RAG1/2 mRNAs occurs specifically in the predominantly cycling, Fraction C' (Figure 1.4A). Therefore, the reduction in RAG expression may also be related to changes in cell cycle distribution in these populations. Hence, in addition to allowing allelic exclusion, it has also been suggested that RAG gene expression is "paused" to allow for a selection checkpoint in which only clonally-expanded precursors which have undergone functional rearrangements of HCs are chosen for further rearrangements at the LC locus. This two wave expression pattern has also been seen by Melchers and co-workers at the protein level for RAG2 (RAG1 protein was not assessed in these studies). However, unlike its mRNA, RAG2's protein is also cell-cycle-regulated throughout lymphocyte development [55].

As with early B cell development, a regulated two-wave expression pattern of RAGs in thymic differentiation has been described by Wilson and co-workers (Figure 1.4B). Based on studies of cell-sorted thymic subpopulations, the first "wave" of RAG expression coincides with CD44+CD25− DN thymocytes undergoing rearrangement at the \( \beta \), \( \gamma \) and \( \delta \) chain loci, and the second with the occurrence of TCR− DP thymocytes undergoing \( \alpha \) chain rearrangement [41]. Between these two waves, RAG mRNA downregulation is observed in the intermediate CD25− DN population. At this stage of development, the TCR-\( \beta \) chain presumably associates with the pre-T\( \alpha \) chain on early thymocytes prior to TCR \( \alpha \) chain rearrangements. The role of the pre-T\( \alpha \) chain appears to be equivalent to that of \( \lambda5 \) in B cell lymphopoiesis, as targeted disruption of the pre-T\( \alpha \) receptor affects development of \( \alpha\beta \) T cells in the same manner as \( \lambda5 \) affects B cell development. The downregulation of RAGs seen in this population would therefore be equivalent to Hardy fraction C' non-cycling cells or Melchers/Rolink
small pre-BII cells. In this context, the presence of the pre-T cell receptor may signal RAG downregulation for similar purposes seen in pre-B cells: allelic exclusion and a "pause" to allow for selection of in-frame rearrangements prior to the onset of α chain rearrangements.

1.1.5.3 Expression of RAG mRNA in precursor cell lines in vitro

In vitro, RAG mRNA expression is also seen in most mouse and human lymphocyte precursor cell lines [15, 53, 56]. Mouse pre-B cell lines from bone marrow or fetal liver can either be established by transformation with AMuLV (Abelson murine leukemia virus) or grown as primary cultures i.e. long-term proliferating untransformed cell lines in the presence of IL-7. In one study, unstable expression of RAGs was reported (1000 fold variations even within subclones of the same cell line), in apparent contrast to other studies that find stable (<10 fold variations) in RAGs between long and short-term cultured lines. The above study also reported reduced RAG mRNA copy numbers/cell compared to in vivo primary tissue samples and sorted subsets [56]. Similarly, in stromal/IL-7 dependent B cell lines, comparable RAG mRNA expression was found in B220+ c-kit+ cells (Melchers pre-BI) shortly after removal from the organ, but gradually downregulate RAG mRNA expression over prolonged time in culture [57]. It has been speculated that loss of expression occurs because persistent RAG expression in cell lines frozen in a developmental stage in which high RAG levels are normally only transiently expressed in vivo, may be detrimental to cellular metabolism, thereby conferring a selective pressure for lower RAG-expressing clones. Supporting this notion is the deleterious effect, including incomplete thymopoiesis, compromised cellular immunity, and increased rates of mortality, in transgenic mice with RAG1/2 under the control of the proximal lck promoter [58, 59]. RAG expression levels in AMuLV-transformed pre-B cell lines may not just be downregulated over time, but may vary depending on the time at which transformation has occurred. Consistent with differential "waves" of RAG expression seen in vivo, within the context of the HC rearrangement program of pre-B cells, it has been found that AMuLV pre-B cell lines with DJ rearrangements have more recombinase activity than do those with V(D)J rearrangements [60, 61].
1.1.5.4 RAG mRNA expression in antigen receptor-expressing cells from primary lymphoid tissues and the third "wave" of RAG mRNAs

Several groups have described in vivo and in vitro situations in which lymphocytes with functional surface antigen receptors can still continue to express RAG mRNAs, albeit at reduced levels ([40, 42, 53, 62]; reviewed in [54]). First, in the T cell lineage, Turka et al. described a population of double positive cortical TCR+ thymocytes that express RAG1/2 [40]. Then, in vitro, Ma et al. described a B cell line established from eμ-N-myc transgenic mice that co-expressed sIgM and the RAGs [62]. In vivo, Alt's group first detected a small population of bone marrow sIg positive B cells (~5% sIgM+ B cells in normal Balb/c mice) that expressed the RAGs by in situ hybridization [62]. Then, Hardy's group detected RAG1/2 mRNA expression in fraction E (sIgM+, immature B cells) by RT-PCR [42]. Melchers and co-workers have also detected low, but detectable levels of RAG1 but not RAG2 mRNAs (the inability to see RAG2 could be due to the sensitivity of the assay) [53]. However, high levels of RAG2 protein were found by these workers in the same immature B cells assessed (RAG1 protein expression was not assessed in these studies). Since Rag2 has a long half-life [63] and is known to be cell cycle regulated [64], one possibility is that these high levels are carry-over from previous differentiation stages.

1.1.5.5 Re-expression of RAG1/2 mRNAs in peripheral lymphoid tissues: the fourth "wave" of RAG expression in GC subsets

RAG expression was thought unique to primary lymphoid tissues and not expressed in SP T cells or mature B cells. That is, after completion of functional H and L chain primary rearrangements (or secondary rearrangement in the case of non-productive or autoreactive specificities) in bone marrow B cells, RAG expression was thought to cease upon migration into the periphery. However, recent literature has also documented RAG expression in peripheral lymphoid organs. For example, high levels of both RAG transcripts have been reported in both gut intraepithelial tissue [65, 66], and adult mouse liver [67]. Both sites are microenvironments that may permit extrathymic T cell maturation. Similarly, Guy-grand et al. demonstrated by in situ hybridization, that 10-20% of peripheral blood lymphocyte (PBL) B cells may express RAG1/2 [65]. Additionally, through RT-PCR [68] or in situ hybridization [69], very faint but detectable RAG expression has been detected in normal mouse lymph
nodes and spleen. RAG1/2 transcripts have also been reported as present in the appendix of young rabbits, a peripheral tissue analogous to the chicken bursa where B cell diversification can occur [70].

In many of the above examples, however, it was unclear whether this mostly low "RAG expression" was biologically significant or was due to contaminating RAG+ immature lymphocytes present in a heterogeneous population. However, recently, significant levels of RAG1/2 expression has been consistently found predominantly in two GC populations: 1) activated mature mouse B cells [71-73] and 2) peripheral CD4+ SP T cells [74]. Furthermore, I have previously shown that the human IgM+ mature B lymphoma cell line OCI LY8, and its variants (which represent in vitro GC counterparts), express high levels of RAG1/2 [75]. As with the signals and cellular mechanisms involved in the re-induction/downregulation of RAGs in bone marrow, a detailed discussion of those involved in the differential regulation of RAGs in the periphery will be included in section 1.2.

1.1.5.5.1 GC formation and architecture

In order to understand more precisely the expression and potential role of RAG products in GCs, the formation and general architecture of GCs themselves must be reviewed (Figure 1.5). GCs are dynamic microenvironments of B cell differentiation that are formed transiently following antigenic stimulation in the B cell follicles of secondary lymphoid tissues such as lymph nodes, Peyer's patches, tonsils, and the per arteriolar lymphocytic sheaths (PALS) of the spleen (reviewed in [76, 77]). In brief, cells that have successfully made functional Ig HC and LC rearrangements exit the bone marrow via the central venous sinusoid, peripheralize and are stimulated by appropriate antigens in secondary lymphoid organs. In the spleen, the white pulp composed of lymphoid cells is organized around a central arteriole. This area is divided into PALS and a surrounding B cell-rich primary lymphoid follicle. Antigen-presenting cells, including interdigitating follicular dendritic cells (FDC) in PALS, entrap the antigen and present them to T_H cells, thereby leading to the clonal expansion of the antigen-specific T cells. The interaction of B cells with activated T cells in outer PALS results in the proliferation of antigen-specific B cells and further T cell expansion. These activated T and B cells in association with FDCs migrate into the primary follicle and form an aggregated structure termed a GC. GCs are compartmentalized into a dark zone containing rapidly proliferating sIg+ centroblasts and a
Figure 1.5. Structure of a typical GC and model of RAG expression and function in GC subsets. A. Typical GC reaction and architecture. B. Putative role for the RAG proteins following RAG expression. B cell clones in GCs which may be rendered autoreactive after an initial round of somatic hypermutation and selection, may re-express RAG, revise their antigen receptors and thereby avoid apoptosis. A legend is shown at the bottom indicating the various cells involved in these processes. Figures are based on those of Pulendran et al., 1997, Ohmori and Hikida, 1998 and Kelsoe, 1998 (references [76, 77, 428]).
light zone filled with sIg* centroblasts, FDC, and CD4+ T cells.

Centroblasts undergo hypermutation in Ig variable regions at a high frequency during rapid proliferation in the dark zone and then migrate to the basal light zone where they differentiate into nondividing, sIg* centrocytes. A portion of these centrocytes is thought to return to the dark zone and re-enter the proliferation cycle. B cell clones diversified by somatic hypermutation are selected through the interaction with FDC, which present immune complexes. B cell clones with higher affinity are positively selected and differentiate into plasma or memory cells in the apical light zone [78]. Conversely, low-affinity, non-productive, or autoreactive clones may be selected negatively, either via deletion or anergy. The selection for B cell clones with higher average affinities is referred to as "affinity maturation".

The interaction between centrocytes and antigen-specific CD4+ T cells mediated by CD40/CD40 ligand and Th2 cytokines (including IL-4) results in isotype switching of secreted antibodies (reviewed in [79, 80]). Furthermore, exposure to different lymphokines or cellular interactions in the apical light zone is believed to determine if the B cell subsequently differentiates into either a memory or plasma cell. For example, interaction of centrocytes with CD40 ligand on T cells alone ex vivo can result in memory B cell differentiation while exposure of the same centrocytes to CD23 expressed on dendritic cells and IL-1α can result in plasma cell differentiation [81]. The micro-environmental architecture of a typical GC is shown in Fig.1.5B.

1.1.5.5.2 RAG expression in GC subsets

All reports in mature mouse B cells have shown that the "fourth wave" of RAG re-expression and accompanying V(D)J double strand break (DSB) intermediates occur in a light-zone subset (PNA*, GL-7*) of mouse mature GC B cells [71, 73]. Although there is currently no precise subfractionation scheme in place for defining mouse GC subsets, these are believed to correspond to centrocytes. In human, a more elaborate subfractionation scheme exists. Pascual et al. were able to distinguish five different subsets of human tonsillar B cells, termed Bmj-Bms, based on reactivity with several different monoclonal antibodies [24]. Cells were initially sorted (by two-color immunofluorescence)
for expression of IgD and CD38. The IgD+ cells, which also expressed IgM were further sorted into CD23- or CD23+ populations (Bm1 and Bm2 respectively). The IgD- population that were CD38+ (CD38 is a marker expressed on GC B cells) were further divided into CD77+ or CD77- (CD77 is a marker that separates dark zone centroblasts from light zone centrocytes). The cells that were IgD- and CD38 were termed Bm5 and represent memory B cells. Thus three different subsets representing different developmental stages were defined; IgM+,IgD+ stage (Bm1 and Bm2), GC stage (Bm3 and Bm4) and memory stage (Bm5). Of interest, Pascual et al. showed that hypermutation had only occurred in the Bm3-Bm5 subsets. Consistent with mouse studies, using the aforementioned human GC subfractionation criteria, a recent study has shown that RAG is expressed at levels comparable to bone marrow B cells in centrocyte subpopulations (Bm4 cells; IgD+, CD38+, CD77) [82]. In striking contrast, little or no RAG expression could be seen in centroblast or follicular mantle populations. The expression pattern of RAG in a typical GC is shown in Fig.1.5B.

1.1.5.6 The RAG-deficient background in lymphocyte development

Homoygous deletions of either RAG1 or RAG2 in mice results in the inability of developing lymphocytes to initiate V(D)J recombination, leading to a complete lack of mature T and B cells [50, 83]. However, in RAG1+ or RAG2+ mice, the early arrest in lymphocyte development is not as complete as the arrest in V(D)J rearrangement, since both committed B and T lymphocyte progenitors are present at elevated levels compared to normal mice. With respect to B cell development, RAG+ mice are blocked at Hardy's fraction C (Figure 1.6A and [50]) or to Melchers/Osmond pro-B/pre-BI cells (Figure 1.6A and [83]). The stage at which B cell differentiation is blocked in mice with targeted deletions in the λ5, the JH locus, or the IgH TM region is strikingly similar to the RAG disruption, but differs in the sense that the latter knockout mice have impaired, but incomplete blocks in V(D)J rearrangements [84-86]. In the context of T cell differentiation, RAG+ mice are blocked at the Thy-1+, CD25+, CD4+CD8- double negative (DN) stage in the thymus (Figure 1.6B) [50, 83]. Analogous to the pre-BCR in B cell development, disruptions in pre-TCR components (β chain, CD3e, or pTα) results in a block in thymic differentiation at the same stage as the RAG disruption, but in only a partial block in V(D)J rearrangement events.
Figure 1.6. The RAG disruption in mouse lymphocyte development. Analogous cells in B and T cell development are represented by the same shading scheme. A. The effect of the RAG disruption in normal mouse B cell development. The stage at which B cell differentiation is blocked in RAG deficient mice is diagramed with respect to FACS analysis for surface expression of normal mouse B cell differentiation markers B220, CD43, BP-1 and HSA and the pre-BCR/BCR complexes. Developmental progression rescued by HC and LC transgenes is represented by dashed arrows. The particular subfractionation scheme and associated letter nomenclature is derived from Hardy (reference [22]). B. The effect of the RAG disruption in the progression of normal mouse thymic development. The block in differentiation caused by the absence of RAG and subsequent restoration with TCR polypeptide transgenes is diagramed with respect to the differentiation of TCR αβ lineage thymocytes in a normal Balb/c mouse. The classification of thymic subpopulations is based on the scheme proposed by Godfrey and Zlotnik, 1993 and Wilson et al., 1994 (reference [41]).
The introduction of functionally assembled antigen-receptor genes or other potentially relevant genes into the RAG\(^{+}\) background can circumvent the V(D)J rearrangement block and allow differentiation to proceed [36]. Specifically, introduction of a mouse \(\mu+\delta\) transgene into the RAG2\(^{+}\) background or a human \(\mu\) transgene into the RAG1\(^{+}\) background promotes B cell development to the Hardy fraction D stage (small, B220\(^{+}\), CD43\(^{-}\) bone marrow pre-B cells) [87, 88]. This observation is also consistent with the previous finding that the expression of a \(\mu\) HC is required for the pro-B to pre-B cell transition [89]. In contrast, the introduction of a functionally rearranged Ig LC gene (\(\kappa\) or \(\lambda\)) does not have any effect on B cell differentiation. Introduction of both rearranged HC and LC transgenes leads to the generation of immunocompetent sIg\(^{+}\) B cells that migrate to the periphery and populate secondary lymphoid organs in relatively normal numbers. Analogous to the rescue of B cell maturation by the introduction of a Ig HC transgene, introduction of a TCR \(\beta\) transgene promotes thymocyte differentiation to a Thy-1\(^{+}\), IL-2R\(^{-}\), CD4\(^{+}\)CD8\(^{+}\) double positive (DP) stage and increases thymic cell numbers to normal levels (100 fold) [90, 91]. Introduction of a functional TCR \(\alpha\) transgene alone does not promote any thymic differentiation or changes in thymic cellularity. However, introduction of both TCR \(\beta\) and \(\alpha\) transgenes leads to the further differentiation of DP thymocytes to CD4\(^{+}\) or CD8\(^{+}\) single positive (SP) T cells that populate the thymus and periphery in relatively normal numbers.

Other signaling-mediating factors can also restore lymphocyte differentiation in RAG\(^{+}\) mice. For example, expression of an activated form of the \(lck\) gene (a signaling intermediate important in TCR signal transduction, \(\beta\) chain allelic exclusion, \(\alpha\) chain rearrangement and normal T cell development progression), in RAG\(^{+}\) mice can restore DP thymocyte numbers to normal levels [92]. Additionally, early thymic signaling structures and their potential role in progression of early T cell development has been demonstrated. In particular, anti-CD3e treatment of RAG1 or RAG2\(^{+}\) mice (which do not express \(\beta\) chain), similar to \(\beta\) or activated \(lck\) transgenes, can also induce DP cell differentiation [93, 94], suggesting that a pre-TCR-like complex is expressed at low, undetectable levels on the surface. Similarly, functional studies in \(\mu\) transgene-complemented RAG\(^{+}\) mice have indirectly supported this possibility by demonstrating the induction of tyrosine phosphorylation in mice following surface engagement of \(\mu\) protein [88]. Taken together, these observations suggest that RAG1/2 are not
required for the commitment to B or T cell lineages or in regulating early lymphocyte development. Rather, it appears that their ability to confer V(D)J rearrangement subsequently allows differentiation to proceed. However, signaling-independent factors may also restore lymphocyte differentiation in mice. In particular, it has been reported that sublethal doses of γ-irradiation can promote a small fraction of pro-B fraction C cells to the small pre-B, fraction D stage [95], although a similar study found that γ-irradiation promoted T cell differentiation only, to the DP stage [96].

1.1.6 Classes of genes co-expressed with RAG mRNAs in antigen-independent and/or antigen-dependent phases of B cell development

The abundance of phenotypic and molecular markers that allow the specific stages of B cell development to be defined is an attractive feature of studying B lymphopoiesis. Since these markers are regulated in a systematic manner, they are a potent method for the characterization of cells and molecular mechanisms involved in antigen-dependent and-independent B cell development. A major aim of this thesis is to better understand developmental events in B lineage in which RAG is expressed. I will therefore review the key categories of genes that are co-expressed with RAGs mRNAs at various stages of B cell development. These can be subdivided into several possible categories: pre-BCR/BCR components and co-receptors, B cell differentiation markers, genes involved V(D)J recombination, or genes involved in the regulation of the RAGs themselves. Table 1.1 and Figure 1.3 summarize the mRNA and protein expression of characterized B-cell and/or stage-specific markers, including those co-expressed with the RAGs at early and late stages of B cell development. The expression scheme is based on the Hardy mouse nomenclature [22, 23] and other differentiation schemes are shown relative to this.

1.1.6.1 Pre-BCR and BCR components: VpreB, λ5, Igα, Igβ

The SLC components, λ5 and VpreB, are a pair of co-expressed B-lineage and stage-specific factors whose mRNAs are expressed co-ordinately with the 1st wave of RAG mRNA expression (Melchers/Rolink pro/preBL and large pre-BII cells [16] or Hardy's B220⁺CD43⁺ pro-B cells [42]), but recently have also been shown to be re-expressed with the 4th wave of RAG mRNA expression in mature GC B cells upon LPS+IL-4, IL-7, or CD40 co-stimulation [71] (Figure 1.3, Table 1.1). The
22 kDa λ5 (called 14.1/16.1 in humans) and 16kDa VpreB polypeptides are both found on the λ1 LC locus, and have high homology to Ig genes, the former in its C-terminus to the C and J regions of λ1 LC [97] and the latter in its N-terminus to HC and LC V regions [98]. Unlike Ig genes, neither undergo somatic hypermutation or V(D)J rearrangement (reviewed in [99]). The SLC components can be expressed on the surface of mouse and human pro-B/pre-B cell lines in various ways: at low levels in conjunction with μ HC to form a pre-B cell receptor (pre-BCR; reviewed in [99]), or at reportedly higher levels prior to the production of a functional μ HC in Hardy fraction A pro-B cells either with a "surrogate HC" consisting of undefined "early" proteins termed gp35-65 and gp130, or in mice, with a product of a DJ-rearranged HC (the Dμ protein; [100]). Cross-linking of these various SLC complexes with anti-μ or anti-SLC MAbs results in calcium mobilization and tyrosine phosphorylation, indicating their ability to transduce BCR-independent developmental signals from the environment [101-104]. Such signals are thought required for survival, lineage progression, promoting Ig HC rearrangements, and establishing IgH allelic exclusion in pro-B/pre-B cells. Mice with targeted disruptions in λ5 show effects on B cell differentiation at the B220^CD43^ pre-B stage similar to that observed in RAG^-^ mice (Figure 1.7; [86]). Overall, analysis of λ5^-^ mice demonstrate its critical role in mediating several key B cell developmental processes, including pre-BII cell proliferation [18], counterselection against RFI DJ joints [99], and HC allelic exclusion [105].

The components important for conferring signaling capabilities to BCR and pre-BCR complexes, the type I transmembrane (TM) glycoproteins Igα and Igβ, (also referred to as CD79a and CD79b and encoded by the mb-1 and B29 genes, respectively), are thought to be exclusively expressed on B lineage cells, although Igβ has been recently reported to also be expressed at certain stages of T cell differentiation [106]. However, unlike the SLC components, they are expressed throughout the various B cell differentiation stages, excepting terminally differentiated plasma cells (Table 1.1; reviewed by [19, 107]). The intracellular domains of Igα and Igβ contain a single copy of the immunoreceptor tyrosine-based activation motif (ITAM; [108, 109]; reviewed in [19]). This motif is found in components of other multimeric signal-transducing receptors, such as the CD3 polypeptides γ, δ, and ε, and the TCR accessory chains ζ and η. The association of BCR with Igα and Igβ is critical for its surface expression, since it is retained intracellularly when transfected alone into
Figure 1.7. Mutant phenotypes of signaling components and transcription factors involved in B cell differentiation. Arrows denote the approximate expression pattern of various transcription factors (relative levels are not shown here; for more detail see Table 1.1), and various signs indicate the stages at which defects have been reported in gene-disrupted mutant mice. Stop signs indicate a complete block in differentiation, yield signs denote a partial block in differentiation (i.e. some mature B cells present in the periphery), and yellow arrows indicate presence of a relatively normal size mature B cell population, but a defect in peripheral B cell function, such as mitogen unresponsiveness, antibody secretion, or germinal centre formation. For comparison, the phenotypes of mutant mice with targeted disruptions in various pre-B and BCR components as well as the RAG genes is also shown. The nomenclature of Hardy is used for representing B cell differentiation stages, as gene knockouts have been reported predominantly in this context. The figure is based on Hagman and Grosschedl, 1994 and Henderson and Calame, 1998; references [18, 436].
nonlymphoid or plasmacytoma cells (which do not synthesize Iga), but not upon Iga co-transfection [110]. The function of the Iga and Igβ in pro-B/pre-B cells is less clear, but it is known that like the BCR, various SLC complexes, including both the Dμ-λ5/VpreB complex on pro-B cells and the μ-λ5/VpreB complex on pre-B cells are also required to associate with Iga and Igβ heterodimers for their surface expression and signaling capabilities [104, 111]. In the case of Igβ, mice with a targeted ablation of the entire molecule have a complete developmental block at the pro-B cell stage (Figure 1.7; [112]). Although their D-J rearrangements remain normal, their V-DJ or V-J rearrangements are severely compromised.

1.1.6.2 BCR co-receptors: CD19, CD20, CD21, CD22, and CD45

Several key B cell-specific receptors with important B cell developmental and signaling roles are expressed along with RAGs in early and late B cell development, and as such may be involved in RAG regulation. These are CD19, CD20, CD21, CD22, and CD45. CD19 is a 95 kDa Ig superfamily type I TM glycoprotein with a long, 240 aa cytoplasmic domain (reviewed in [113]). Expression of CD19 is restricted to B lineage cells, beginning in the earliest B-committed progenitors, continuing through all four waves of RAG expression, and terminating at the terminally-differentiated plasma cell stage (Figures 1.2 and 1.3, Table 1.1). Recent studies have identified CD19 as a key signaling molecule at several B cell differentiation stages (reviewed in [114]). Whereas associated molecules and/or counter-receptors for CD19 are not well-characterized in pro/pre-B cells, in immature and mature B cells in vivo, CD19 is functionally and physically associated with the BCR. In mature B cells, also forms a signal transduction complex with TAPA-1 (CD81), the complement receptor CR2 (CD21), and associates with CD77, an uncharacterized GC B cell differentiation marker. Furthermore, CD19 is phosphorylated following α-CD19 crosslinking, as well as following the crosslinking of other receptors; BCR, CD40, and CD72. This suggests a signaling pathway that intersects with other B cell signaling pathways (reviewed in [11]). CD19 has a role in lowering the BCR signaling threshold, as indicated by the observation that CD19+BCR co-ligation enables cells to proliferate in response to much lower anti-μ concentrations than to BCR stimulation alone [115]. However, CD19 ligation renders B cells refractory to subsequent BCR stimulation. Targeted disruption of the CD19 gene in
mice show severely reduced antibody responses to T-cell dependent antigens, further illustrating the important role of CD19 in mature B cells in vivo [116, 117].

CD21 is a 145 kDa protein that serves as a complement receptor (CR2) for the complement cleavage fragments C3d and C3dg [118]. CD21 also serves as a receptor for Epstein-Barr virus (EBV), an herpes simplex virus (HSV) family member that can transform human B cells in vitro and is associated with malignant B cell lymphomas [119]. CD21 is expressed on mature B cells and is also expressed on FDCs, a subset of thymocytes, and on epithelial cells (reviewed in [120]). By association with CD19, CD21 may be involved in B cell activation and in governing signaling thresholds with relevance to autoimmunity [121]; reviewed in [122]). CD21 may also be vital for binding of immune complexes to FDCs. Although CD21 expression is undetectable by flow cytometry on pre-B cells, the fact that EBV can transform pro-B/pre-B cells [123] coupled with the observation that a MAb that prevents binding to CD21 also abrogates EBV transformation of pro-B and pre-B cells indicates the presence of CD21 expression in earlier B cell subsets at levels too low to be detected by conventional methods.

CD22 is an Ig superfamily B lineage-restricted phosphoglycoprotein heterodimer consisting of 130 and 140 kDa chains. Its expression pattern during B cell development provides an example where mRNA and surface protein expression do not correlate (reviewed in [124]). In human and mouse, CD22 mRNAs are found co-expressed through all four RAG expression waves, however CD22 protein, though found in the cytoplasm of pro-B, pre-B and B cells, only appears on the cell surface in a subset of mature B cells, and possibly in mouse on some pre-B cells (Figure 1.3; [125]; reviewed in [126]). This presents one limitation to the interpretation of mRNA RAG co-expression patterns. It also suggests that there is a mature B cell-specific CD22-associated protein, the synthesis of which is required for CD22 surface expression. The CD22+ B cell subset undergoes a Ca²⁺ flux and proliferates in response to anti-µ. CD22 is physically associated with the BCR and has a role in "negative" instead of co-stimulatory or "positive" B cell-signal transduction [127]. In particular, the CD22 cytoplasmic tail contains a sequence motif opposite to the aforementioned ITAM motif that confers this inhibitory signaling and is shared by various other negative signaling receptors, the so-called Immunoreceptor Tyrosine Inhibitory Motif (ITIM) [127]. The ITIM enables the interaction with SH2-
containing domains of cytoplasmic proteins involved in ITAM motif dephosphorylation. One such protein is Shp1, a phosphotyrosine phosphatase known to negatively regulate leukocyte activation (reviewed in [128]), which has its phosphatase activity perpetuated by binding phosphorylated CD22 [129]. A model for CD22 function in B cell signaling has been proposed in which upon BCR engagement, B cells become activated leading to CD22 tyrosine phosphorylation [124]. CD22 then recruits the phosphotyrosine phosphatase Shp1 to the BCR complex, permitting the dephosphorylation of appropriate substrates, thus terminating the activation signal. The hypothesis that CD22 officiates as a negative regulator has been supported by recent studies of CD22−/− mice whose B cells demonstrate hyperresponsiveness [130, 131].

Originally termed leukocyte common antigen, CD45 is a highly-glycosylated 180-240 kDa type I TM phosphotyrosine phosphatase that is abundantly expressed on the surface of all hemopoietic cells except erythrocytes [132, 133]. Multiple isoforms of CD45 exist and are differentially expressed on different cell types and at distinct stages of differentiation and activation [134]. The mRNA encoding the B cell isoform, CD45R (B220 in mice), regulates BCR signaling by dephosphorylating src family PTKs such as lyn as well as Igα and Igβ (reviewed in [114, 135]). Like CD19, B220 is detected on B cells at all developmental stages except terminally differentiated plasma cells. Mice with a targeted deletion of CD45RO exon 6 exhibit a profound block in early T cell development and a defect in later B cell differentiation [136]. Mature B cells, isolated from B220−/− mice proliferate in response to lipopolysacharride (LPS) stimulation, but not BCR ligation, suggesting CD45’s crucial role in regulating the physiological response to foreign antigens via BCR aggregation.

1.1.6.3 Stage-specific cell-surface differentiation markers

One functionally and structurally diverse subclass of genes co-expressed primarily with RAG in early B cell development are those encoding cell-surface differentiation markers. Although the CD antigens CD10, BP-1, and CD43, and CD34 are not B lineage-specific, within the B lineage, their mRNAs are temporally related to the first two waves of RAG expression at antigen-independent phases of B cell differentiation (Figures 1.2, 1.3, and Table 1.1). In addition, like RAGs, some of these markers (such as CD10 and CD43) are re-expressed on activated peripheral B cells during the antigen-
dependent phase of B differentiation. Selected subsets of mature B cells engaged in modification of their antigen receptor genes, in particular GC centroblasts and centrocytes are positive for CD10 and CD43 (reviewed in [22, 137]; Figure 1.3 and Table 1.1). CD10 (also called CALLA) is a 80 kDa type I human-specific, cell surface, endopeptidase ectoenzyme belonging to the metallopeptidase family (reviewed in [138]). BP-1/6C3, which has a similar structure and expression pattern as CD10, is a 110-150 kDa type II TM antigen with an extracellular aminopeptidaseA activity (further detailed in section 1.5 and reviewed in [139]). Both BP-1 and CD10 may have roles in lymphopoeisis by cleaving regulatory peptides in the hematopoietic microenvironment. CD43 (Leukosialin) is a 90-120 kDa type I TM protein with a highly O-glycosylated extracellular domain, and long intracellular domain (reviewed in [140 a major]). The primary role of CD43 is in cell proliferation and cell adhesion. Mice overexpressing CD43 have decreased susceptibility to B cell apoptosis and increased numbers of B cells [141]. Although not the primary cause, defective expression of CD43 has also been associated with the X-linked immunodeficiency Wiskott-Aldrich syndrome in humans [142]. Finally, CD34, which may function as a pro-B cell adhesion molecule, is not essential for hematopoiesis, since its targeted disruption does not significantly alter mouse blood cell or lymphocyte development [143].

1.1.6.4 Co-expressed and/or associated genes involved in V(D)J recombination

1.1.6.4.1 Terminal deoxynucleotidyl transferase (TdT)

The best example of a gene co-expressed with RAGs and whose product is involved in the V(D)J recombination process is TdT, which encodes a lymphoid specific, developmentally regulated DNA polymerase. TdT mRNAs are expressed only in the first wave of mouse RAG expression, and show a very similar expression pattern to λ5 and VpreB mRNAs (Figure 1.3, Table 1.1; [42, 144]). Besides the Rags, TdT is the only other recombinase component identified thus far with a characterized function in the V(D)J process that is also lymphoid-specific. In humans, as already mentioned, TdT mRNA expression differs from that in mouse B differentiation, in that they are also expressed in the second wave of RAG expression, during additional LC N-additions. TdT can add nucleotides to the 3'-OH end of DNA in the absence of a template in vitro. Its physiological role is in the addition of N
additions at Ig and TCR coding region junctions, thereby increasing sequence diversity [145].

The evidence supporting the role of TdT in N addition comes from several observations. First, TdT adds deoxynucleotides in a non-templated fashion in vitro with a preference for adding G residues, consistent with the GC-rich sequence found in most N regions in vivo [146]. Secondly, TdT is expressed in early B and T cells of adult mice, but is not expressed during fetal development of most vertebrate species, excepting in the human [147]. This corresponds to the absence of N additions in antigen receptors rearranged during fetal ontogeny [148]. Furthermore, in mouse B cell development, TdT expression is highest in pro-B cells undergoing Ig HC rearrangements i.e. co-expressed with the first wave of RAG expression, but is terminated in pro-B/pre-B cells undergoing LC rearrangements [15]. This is consistent with the observation of a low incidence of N region diversity in the mouse LC repertoire [149]. In human, TdT and N-additions are both found in pre-B cells undergoing LC rearrangement (co-expressed with the first two waves of RAG expression). Thirdly, the incidence of N regions in a pre-B cell line is increased by expression of the cloned TdT gene [150]. Further evidence comes from the analysis of recombinants made in fibroblasts expressing RAG1/2 [151]. These cells do not produce TdT, and coding junctions are devoid of random base insertions. Co-expression of TdT then results in coding joints with inserts like those found in lymphoid cells. The proposition that TdT introduces N regions has been definitively confirmed by the observation that very few N additions in Ig and TCR gene rearrangements are found in lymphocytes of mice in which the TdT gene has been disrupted by homologous recombination [152, 153]. In these mice, recombination junctions routinely occur at regions of homology resulting in a severe restriction in the repertoire (the repertoire of an adult TdT- mouse closely resembles that of a neonate). However, TdT- mice are capable of mounting efficient immune responses, thus casting doubt of the necessity of N addition for normal B and T cell function. Therefore, unlike the RAG products, participation of TdT in both V(D)J recombination and lymphocyte development is optional.

### 1.1.6.4.2 Potential, non-essential components

The Rags are the only known lymphoid-specific and required factors for V(D)J recombination. All other essential recombinase components are ubiquitously-expressed DNA double strand break repair
Prior to the molecular identification of these DSBR components, efforts to find factors essential for V(D)J rearrangement focused on predicted biochemical activities involved in V(D)J recombination such as RSS binding, cutting, and ligation (Table 1.2). As a result, nearly a dozen factors were identified and reported to recognize (bind) or cleave the heptamer, the nonamer, or both although none had been linked convincingly to V(D)J recombination. Of these, a few had lymphoid-restricted expression patterns (Table 1.2). The first to be reported was called "nonamer-binding protein" (NBP) and was identified on its ability to complex stably with a 23-spacer RSS [154a]. After further purification and characterization, NBP, a 53 kDa protein found to be expressed only in lymphoid nuclear extracts, was demonstrated to be very specific for the nonamer sequence, but did not possess any nucleolytic activity. A fourth protein with an apparently early B cell-specific expression pattern was isolated as a cDNA clone by screening a pre-B cell library with a 12-spacer signal probe [155]. This clone, T-160, was specific for 12-spacer RSS and failed to bind a sequence with a mutation in the heptamer's third position. The role of the T-160 protein in V(D)J joining was questionable because its binding properties as determined by Southwestern blot analysis could not be reproduced in EMSA or DNA footprinting assays. Perhaps the most promising candidate among those identified on the basis of binding is a 30 kDa protein, called "recognition protein" or RP, found in mouse thymocytes [156]. RP is specific for joining signals while binding to both heptamer and nonamer targets in both 12 and 23 RSS; it binds non-functional joining signals poorly, and is confined to cells and tissues that exhibit recombination activity. With respect to ligation activity, a 47 kDa product was isolated based on its nonamer-binding properties [157]. This clone, called "V(D)J joining protein" (VDJP) has a domain homologous to bacterial ligases and when expressed bacterially, had a joining signal-dependent activity that was greatly reduced upon deletion of either RSS heptamer. This activity also correlated with pre-B cell line expression, as did lack of activity in fibroblasts. However, despite this apparent specificity, the ligation products did not resemble signal junctions.

The recent finding that a complete cell-free V(D)J reaction can be reconstituted with little else except the Rags+DNA ligases [158-160] suggests that if any of the aforementioned proteins have a role in V(D)J recombination, at least in vitro, they are likely to act only as modulators of recombination efficiency (for example, by competing with Rags for signal binding). In my estimation, it is not likely that that
<table>
<thead>
<tr>
<th>Protein</th>
<th>Trans-acting components involved</th>
<th>V(DJ) recombination:</th>
<th>Recombination activating gene 1 (RAG1)</th>
<th>Recombination activating gene 2 (RAG2)</th>
<th>Non-essential, lymphoid specific</th>
<th>Non-essential, myeloid specific</th>
<th>Essential</th>
<th>Essential</th>
<th>Regulating protein</th>
<th>Regulating protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rad51</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>119,000</td>
<td>54,000</td>
<td>none</td>
<td>none</td>
<td>54,000</td>
<td>54,000</td>
<td>54,000</td>
<td>54,000</td>
</tr>
<tr>
<td>XRCC4</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>38,000</td>
<td>38,000</td>
<td>none</td>
<td>none</td>
<td>38,000</td>
<td>38,000</td>
<td>38,000</td>
<td>38,000</td>
</tr>
<tr>
<td>KU86</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>Rad52</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>38,000</td>
<td>38,000</td>
<td>none</td>
<td>none</td>
<td>38,000</td>
<td>38,000</td>
<td>38,000</td>
<td>38,000</td>
</tr>
<tr>
<td>ATR</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>ATM</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>53BP1</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>Cdc14</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>WACS</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>53BP1</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>MDC1</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>TANK-binding protein</td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>NBS1</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>NBS1</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>NBS1</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>NBS1</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>NBS1</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
<tr>
<td>NBS1</td>
<td></td>
<td>DNA double-strand break repair</td>
<td>78,000</td>
<td>78,000</td>
<td>none</td>
<td>none</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
<td>78,000</td>
</tr>
</tbody>
</table>

**Table 12:** Table of Trans-acting components involved in V(DJ) recombinations, the endonuclease enzymatic machinery.
the DD RT-PCR screens would reveal any other essential, lymphoid-specific recombinase components (besides the Rags themselves). Therefore, the details of V(D)J biochemistry will not be discussed further in this thesis; the reader is consulted to several excellent reviews which have been recently been published, both on the essential factors involved at various stages of V(D)J recombination [161] as well as on the plethora of recent advances which have contributed to knowledge of the biochemical mechanisms involved in this process [162-164].

1.1.6.4.3 Rag-associated proteins

Consistent with the close functional co-operation of their products in the context of extrachromosomal V(D)J recombination assays and cell-free cleavage reactions, direct biochemical evidence shows that the RAG products are physically associated. Furthermore, the Rags are also associated with other recombinase proteins in vivo and in vitro [165, 166, 167]. The genes encoding such products is potentially another class of genes that may be identified in my screens. Although not necessarily factors that may be co-expressed with RAG, it is therefore noteworthy to mention other genes whose products physically associate with the RAG proteins and that have been identified based on this association. In particular, there is indirect evidence that the aforementioned recombinase component TdT also associates with the Rags, thus providing an example of a RAG co-expressed gene whose product also may physically associate with the Rags (described below). Other more ubiquitous proteins, such HMG, and the DSBR DNA-dependent protein Kinase (DNA-PK) subcomponents associate with Rags at different stages of the V(D)J reaction (reviewed in [164]). While the biochemical function of these proteins at various stages of the reaction will not be discussed, I will touch on their potential roles with respect to their association with Rags and other proteins in Rag complexes. In order to discuss the associations of Rags with other proteins, the structure of the Rags and characteristics of their physical association with each other will first be reviewed.

RAG1/RAG2 RSS-dependent and independent complexes

Using co-immunoprecipitation experiments, Leu and Schatz first demonstrated that Rag1/2 are components of the same complex [168]. These experiments also revealed that three distinct populations of in vivo Rag1 molecules may exist: one population bound to multiple Rag2 molecules;
another Rag1 population that is not in these complexes is tightly bound to RNA or DNA, and a Rag1 dimer population. Since α-Rag antibodies were equally effective at co-immunoprecipitation with or without nucleic acid removal in these assays, this suggested RSS independent-Rag1/2 associations. Schatz and co-workers then used Rag1/2 deletion mutants in co-immunoprecipitation assays to determine two regions of Rag1 responsible for interacting with Rag2 in vivo: aa 504-570 and 850-1008, both in the recombinationally active rag1 "core" (see Figure 1.8; [169]). However, this study still did not address the issue of whether Rag1/2 interactions are direct or indirect.

In apparent contrast to the above studies of full-length Rags in vivo, the "core" Rags in vitro were initially reported capable of associating with each other only in the presence of a RSS [170]. However, recently, the "core" Rag1 has been found to homodimerize in a DNA-independent fashion [171]. The authors suggest that this is a transient interaction, detectable only by sensitive methods such as glutaraldehyde crosslinking, and may explain why only more stable RSS/Rag1/2 complexes could be detected previously. Using yeast two-hybrid assays, the Rag1/2 interaction has also been incapable of being reproduced in three independent reports [165, 166, 172]. This indicates either one or more, proteins mediate Rag1/2 complex formation or important post-translational modification of Rags fail to take place in yeast cells, (the latter which would explain the lack of Rag1/2 association in yeast two-hybrid assays).

Rag-interacting proteins Rch1 and SRP1

In search of candidate proteins that interact with the RAG1 and RAG2 products (Rag1 and Rag2, respectively) using a yeast two-hybrid assay, two related proteins that specifically associate with Rag1, Rch1 (RAG cohort-1) and the human homologue of SRP1 (suppressor of temperature-sensitive RNA polymerase I mutation), have been identified [165, 166]. Rch1 and SRP1 are ubiquitous 58 kDa nuclear products which bind via four repeats to basic-charged binary N- and C-terminal regions of Rag1 [172]. Co-transfection of Rag1 with deletional mutant forms of Rch1/SRP1 does not significantly affect recombinase activity compared to those with wild type Rch-1, as determined by extrachromosomal recombination assays. Furthermore, the Rch-1-associating region of Rag1 is not required for recombination [173-176]. Therefore, these proteins are not required for in vitro V(D)J
Figure 1.8. Structure of the RAG proteins. The "active core" regions of mouse Rag1 (aa 384-1010) and Rag2 (aa 1-387) i.e. the minimal regions necessary for V(D)J recombination in the context of plasmid substrates expressed in fibroblast cells (and in vitro cleavage reactions), as determined by a number of independent structure/function mutational studies (Cuomo et al., 1994; Sadofsky et al., 1993; Silver et al., 1994; references [172-175]) are represented by grey-shaded areas. A. Structural features of Rag1. The overall structural domains are shown at the top, and below, the N-terminal region of Rag1 that is not necessary in vitro, but may be important in vivo in formation of an active recombination complex (Roman et al., 1996) is shown at a higher resolution to outline the various conserved residues across species within the cysteine/histidine-rich and Zn/RING domains characterized in this region. Regions of Rag1 which interact with the Rag1-associated proteins SRP1 and Rch1 (as determined by yeast two hybrid analysis) are shown (aa 49-273 and 818-1019, respectively). Five basic regions involved in nuclear localization are shown by + signs (aa 141-146, 221-224, 243-249, 826-828/838-840, and 969-973, respectively). Abbreviations for other domains are as follows: Cys-His, Conserved cysteine and histidine-rich region (aa 15-168); dimer; Zinc-RING and Zinc-finger dimerization domains (aa 269-389); GGRPR, region with homology to the GGRPR Hin recombinase homeodomain (aa 389-446); homology to HPRT topoisomerase active site motif (aa 994-998). B. Structural features of Rag2. Abbreviations are as follows: Ser 356; serine 356 phosphorylation site (critical for recombination activity); Thr490; threonine 490 phosphorylation site (critical for cell-cycle regulated expression of rag2); acidic, acidic transcription factor-like domain (aa 382-409).
recombination. However, it remains to be determined if they are required for endogenous V(D)J locus rearrangements. In particular, yeast SRP1 has been shown to bind to the nuclear envelope [174] and the Rag1 Rch1/SRP1 binding regions contain putative NLS (which mediate the nuclear and nucleolar translocation of proteins). Rch1/SRP1 are therefore implicated as having a role in the nuclear localization and/or nuclear transport of Rag1. In this context, it is possible by localizing Rag1 and/or Rag1/2 complexes (and possibly other recombinase components like TdT) to the nuclear envelope, this allows proper synapsing of endogenous antigen receptor segments separated over large distances. The importance of such a nuclear localization effect is observable in other cellular processes such as replication and transcription [177, 178]. Curiously, the yeast two-hybrid systems used to identify the above two Rag1-associated proteins SRP1 and Rch1 failed to detect any Rag2-interacting proteins, including Rag1 itself, making the physiological relevance of the RAG-Rch1/SRP1 interactions currently unclear.

As would be expected for proteins involved in recombination, and consistent with the presence of Rag1 NLS and association with Rch1/SRP1 proteins, immunofluorescent staining of cells with α-Rag antibodies reveal that Rag1/2 are predominantly located in the nucleus in both thymocytes and transfected cell lines [53, 172]. In normal cells, the predominant subcellular localization pattern of both Rags is in the nuclear periphery, suggesting localization to the nuclear membrane [172]. However, in fibroblasts transfected with either Rag1 or cotransfected with Rag1/2, α-Rag1, α-Rag2, and α-fibrilliar staining reveals a nucleolar subcellular localization pattern [168]. In contrast, fibroblasts transfected with Rag2 alone shows α-Rag2 staining in the nuclear periphery, but none in the nucleolus. Therefore Rag2 is not intrinsically associated with the nucleolous, but is localized there in a Rag1-dependent manner. This is consistent with Rag1, but not Rag2 having various NLS as well as regions for Rch1/SRP1 interactions (Figure 1.8). The different subcellular localization of a free Rag2 pool relative to the complexed form may also indicate an additional, unknown role for Rag2 in vivo. Interestingly, like the RAG and Rch-1 proteins, TdT appears to associate with the nuclear matrix, as determined by direct immunofluorescence and cellular subfractionation [179, 180], raising the possibility this factor may physically complex with the RAG1/Rch-1 protein complex.
RAG/HMG1/2 cleavage and post-cleavage complexes

In mammalian cells, the chromatin-associated high mobility group proteins HMG1 and HMG2 are abundant DNA-bending factors that are believed to play an architectural role (including the assembly of nucleoprotein complexes) by bending DNA into conformations which facilitate interactions between other proteins [181]. It has therefore been hypothesized that HMG1/2 complex with the Rags for increasing cleavage efficiency and have a physiological role in RSS synapsis (Table 1.2). Two reports have supported the hypothesis of Rag1/2/HMG1/2 cleavage complexes \textit{in vitro}. First, Gellert's lab has shown that HMGs supplemented to Rags enhance cleavage activity, and are particularly efficient at stimulating coupled cleavage i.e. concerted cleavage of both RSS, suggesting that these proteins also aid in the formation of a synaptic complex [181]. In the other report, HMG1/2 were found to enhance Rag1/2-mediated cleavage of substrates containing both 12 and 23 RSS motifs but not the 12 RSS motif only [182]. Recent \textit{in vitro} and \textit{in vivo} evidence of HMG/Rag1 interactions have provided further insight into the role of Rag-associated HMGs in cleavage [183]. By gel shift analysis, the Rag1 homeodomain was shown to be important for the \textit{in vitro} Rag1-HMG1 interaction, as its deletion resulted in loss of this specific interaction (See Figure 1.8). A role for HMGs in synapsis was suggested by their increasingly pronounced effects when RSS in artificial substrates were spaced over larger distances. Furthermore, \textit{in vivo}, binding to Rags by co-immunoprecipitation of Rag1 with α-HMG1 antibodies and HMG-enhanced cleavage (using luciferase reporter assays) were demonstrated.

Taken together, the above reports argue that only Rags are required to complex for relatively efficient single RSS cleavage \textit{in vitro}. However, for optimal \textit{in vitro} coupled cleavage, especially for substrates with distantly spaced RSS, HMGs are also required. It is therefore likely that for efficient \textit{in vivo} cleavage/synapsis of endogenous antigen receptor loci to occur, the physiological cleavage complex is a large, multi-nucleoprotein amalgamation comprised of Rags (both which are essential for cleavage), TdT, the Rag-associated nuclear-localizing proteins Rch1 and SRP1 (for proper targeting of the Rags), and the DNA-bending proteins HMG1 and 2 (for proper synapsing of RSS spaced over physiological distances).

Several papers have provided direct evidence that both Rags also complex with HMGs and other proteins in post-cleavage activities [167, 184, 185]. Using a coupled cleavage system, both the Schatz
and Gellert labs have shown by a combination of nuclease sensitivity, EMSA, and/or immunoprecipitation assays that a stable complex of Rag1/2, HMG1/2 and the DNA-PK subcomponents (comprised of DNA-PK\(_{\text{ca}}\), the catalytic subunit, and the Ku 70/86 heterodimer) is formed post-cleavage, between synapsed RSS. It could also be hypothesized that the TdT, which co-localizes with RAG in the nuclear membrane and is important for end modifications, may also associate (directly or indirectly) with the Rags in this post-cleavage complex. It is thought that such a complex may be required for bringing all four ends in close proximity to each other for end-modification and/or for resolving coding ends. The former possibility is suggested by the occurrence of nonstandard hybrid joints, which involve the 'mistaken' joining of a coding end to the other signal end instead of to the other coding end, an occurrence that likely requires all four ends remaining in proximity until ligation. Recently, it has been shown that the post-cleavage activities of the Rags include hairpin coding-end opening [186] and resolution of non-standard hybrid joints and possibly normal end-joining [158, 185].

While the HMGs likely have a role in making antigen receptor locus segments more accessible to Rags in cleavage complexes (and possibly for Rags and other recombinase components in post-cleavage complexes), what possible functional significance would the DNA-PK/Rag1/2 associations have with respect to Rag function? One possibility is that the large size of DNA-PK\(_{\text{ca}}\) may provide a framework for other recombinase components, including the Rags to localize/associate with each other and antigen receptor loci. Additionally, the close physical association of DNA-PK\(_{\text{ca}}\) with RAGs may be necessary to phosphorylate the Rags in order activate its hairpin endonuclease and/or end-joining activities. The physical association may also have an indirect functional effect on the Rags, for example, by being in physical proximity with the Ku heterodimer, this activates the latter’s helicase activity, which may be required for unwinding hairpin ends in order to make them accessible for the Rags. Regardless of how this occurs, the association of Rags with other proteins is required for its activity in its post-cleavage functions, as seen by the fact that hairpin endonuclease activity of the Rags was found to be dependent on post-cleavage complex formation [186].
1.1.6.5 Co-expressed genes potentially involved in RAG regulation

Although the molecular and cellular aspects of RAG regulation at each stage of B cell development will be extensively discussed in the next section, here I will introduce a few possible genes that are specifically co-expressed with RAG which also may regulate these genes, based on different lines of direct or indirect evidence. This is an important aspect of this thesis, as my screens for potential RAG regulators are based on RAG mRNA co-expression of such factors. It is also noteworthy that the various classes of RAG mRNA co-expressed genes discussed here may also belong to other classes of RAG co-expressed genes discussed above. For example, any of the BCR or pre-BCR co-receptors could be directly involved in RAG regulation either indirectly via modifying BCR-mediated RAG regulatory signals or directly via their own direct RAG activation pathways.

Cytokine receptors

Evidence indicates that cytokine receptors may themselves regulate RAG expression. One example is the Interleukin-7 receptor (IL-7R). Interestingly, the IL-7R is not only expressed with RAGs in antigen-independent B differentiation, but, like the SLC components and TdT, can also be re-induced co-ordinately with the 4th wave of RAG expression (Figure 1.3, Table 1.1). The IL-7R consists of a chain which incorporates at least two subunits: α and γ. The α chain is shared with thymic stromal cell lymphopoietin (TSLP) and the γ chain is common to the receptors for IL-2, IL-4, IL-7, IL-9 and IL-15 (reviewed in [187]). In mice, IL-7 serves as an essential growth factor for pre-B cells, and IL-7R expressed on pro-B and early pre-B cells has been shown to mediate a proliferative response to IL-7 (reviewed in [188]). Injection of neutralizing antibodies against the IL-7R, or targeted disruption of either IL-7R subunit all results in a similar phenotype: a severe interruption in early B cell development including a dramatic reduction in pro-B cell bone marrow subpopulations and peripheral lymphoid cellularity [189-191]. Furthermore, during in vitro culture of pre-B cells from IL-7Rα⁻ mice, HCs fail to rearrange. Rearrangement is restored when cells are transfected with the IL-7Rα gene [192]. The presence of some peripheral B cells in IL-7⁺ mice indicates a compensating mechanism. As mentioned earlier, IL-7 is not essential for human expansion/differentiation of B cell progenitors, suggesting different biological effects of IL-7 and IL-7R in mouse and human B cell development.
IL-7 signaling appears to differentially regulate RAG in T and B lineages. IL-7 treatment has been reported to sustain expression of the RAG1/2 genes and enhances TCRβ rearrangement in pre-T cells from mouse embryo thymic organ cultures [195]. Because these studies used an intact thymic organ culture system, it was impossible to determine whether the IL-7 effects were direct i.e. driving gene rearrangement during early T cell development or indirect i.e. simply allowing cells to expand while another factor is in fact responsible for RAG induction. Nevertheless, in contrast, an inhibitory effect for IL-7 on RAG expression in mouse pro/pre-B cell development has been shown. In this study, recombinant mouse IL-7 was found to inhibit Ig rearrangement and B cell precursor maturation, and also possibly influence V(D)J recombination by post-transcriptional downregulation of RAG1/2 [196]. In human fetal BM pro-B cells, IL-7 enhances the expression of CD19 and subsequently downregulates the expression of RAG1/2, suggesting a cascade of IL-7 effects [197]. Additionally, the engagement of CD19 by divalent antibodies completely blocks the IL-7 induced downregulation of RAG expression, suggesting that in human pro-B cells, RAG regulation and Ig rearrangement are modulated by both IL-7 and CD19. In another study, pre-B cell lines derived from Eμ-bcl2 transgenic mice demonstrated proliferation in the presence of stromal cells and IL-7, but removal of IL-7 resulted in transient upregulation of RAG1/2 expression followed by a decrease in their expression upon differentiation into IgM+ cells [198]. This suggests that IL-7 may have a biphasic effect on RAG1/2 expression at different stages of mouse antigen-independent B cell development. Finally, in IL-7Rα+ mice, it was shown that RAG1/2 were transcriptionally active in fetal and adult thymocytes [191]. Therefore, IL-7 is not required for RAG expression, although in some cases it may deliver signals to enhance RAG expression/gene rearrangement as well as the viability of developing cells.

The IL-7R is re-expressed with the 4th wave of RAG expression, but in contrast to early B cell development, the IL-7R signaling pathway in later B development will induce mouse RAG expression [199]. Further, in vitro or in vivo activated GC B cells express high levels of IL-7R and blocking of the IL-7R in vivo with α-IL-7R MAb suppresses RAG induction and accompanying secondary λ rearrangements during immunization in draining LN of mice. This IL-7 mediated induction of RAG,
rather than downregulation, suggests that IL-7 acts differently on progenitor and mature B cells. Additionally, IL-7 in the GC microenvironment has also been shown to re-induce other early B cell-specific genes, including HSA, TdT, λ5, and VpreB, suggesting that the GC environment may be similar to an IL-7-responsive BM microenvironment [71]. However, in contrast to effects on RAG induction in early T and B differentiation, it acts as an in vitro cofactor, in conjunction with α-CD40. Overall, studies in mice and human cells show reveal complex effects of IL-7 on RAG expression which may include differential regulation of RAG expression between mouse and human B cells, between mouse proB and proT cells, and even within the same B or T cell progenitors.

**BCR co-receptors**

Since the BCR itself (and possibly also the pre-BCR) has an important role in RAG regulation, it would not be surprising that various of its components may regulate RAGs. First, it is possible that signaling through SLC components VpreB and λ5 may regulate RAG expression in antigen-independent stages of B cell differentiation. In addition, because pre-BCR or BCR co-receptors can modulate various B cell signaling outcomes, they too may be implicated in regulating RAG expression. Such RAG co-expressed BCR co-receptors could also modulate other B cell signaling pathways, such as cytokine-signals. One such example, CD19, regulates RAG itself by modulating a cytokine-mediated RAG-induced signal (discussed above). In particular, CD19 engagement by divalent antibodies completely blocks the IL-7 induced downregulation of RAG expression [197], suggesting that in human pro-B cells, RAG regulation is modulated by antagonistic signals generated through IL-7 and CD19. The CD19-mediated effect appears to be RAG-specific, because it does not affect the accompanying IL-7 mediated downregulation of TdT. A further functional relationship between CD19 and IL-7 signaling is observed in B cell precursors: IL-7 itself can upregulate the expression of CD19 on pro-B cells, suggesting a negative feedback regulatory mechanism [200].
**B cell transcription factors**

There is indirect evidence that several B lineage transcription factors may be involved in RAG regulation during antigen-independent stages of B cell differentiation. This is based on two lines of evidence, first, their expression patterns resemble those of RAGs in early B cell development, and second their knockout phenotype provide evidence that RAGs either are, or may be, targets of these factors. Other B lineage transcription factors expressed later with the 3rd and 4th waves of RAG expression in B cells are likewise candidate RAG regulatory factors at these later stages. I will briefly discuss a few examples of possible activators and repressors in early and late development, in terms of their expression and knockout phenotypes. The possible hierarchy of transcription factors involved in RAG regulation will be deferred until later in Chapter 5 as discussed in the context of work presented in this thesis. A comprehensive list of the various transcription factors important in early and late B cell development (based on knockout phenotypes) and/or with stage-specific and/or B lineage-restricted expression patterns are shown in Table 1.1 and Figure 1.7.

Examples of factors co-expressed in early B cell development that are also potential early B cell-specific RAG activators (based on their expression pattern and knockout phenotypes) include EBF, LEF, and Sox-4. Early B cell factor (EBF), a homodimeric transcription factor was identified based on its binding to and activation of the Igα promoter [201, 202]. Its mRNA and protein expression is restricted to the two waves of RAG in early antigen-independent stages of B cell differentiation (Table 1.1). EBF− mice display a block in B cell differentiation prior to RAG− mice, at the Hardy fraction A early pro-B stage (Figure 1.7; [203]). Examination of pro-B cells from these mice reveals the lack of Igα, Igβ, SLC, and RAG1/2 transcripts and DJ₅ rearrangements, but the presence of IL-7R, TdT and μ sterile transcripts. Sox-4 was originally isolated in a screen for factors binding to the T cell-specific enhancer consensus sequence 5'-AACAAAG-3' [204]. Its mRNAs and proteins are expressed in several non-lymphopoietic sites during murine embryogenesis, but in adult mice are restricted to lymphocyte lineages, and within the B lineage, is co-expressed with the first three waves of RAG expression, in pro-B, pre-B and immature B cells [205]. While hematopoietic cells of other lineages remain intact, targeted disruption of the Sox-4 gene causes severe impairment in B cell development just prior to the RAG disruption in Hardy fraction B cells (Figure 1.7; [206]). The presence of low
levels of mature B cells indicates that this B differentiation block is not comprehensive.

It is noteworthy that RAG co-expression does not alone establish a functional role for a particular transcription factor in B cell development and/or B cell-specific RAG regulation. Some critical factors may not be lymphoid-restricted. An example is the ubiquitously-expressed E2A gene, which is critical not only for Ig transcription and B cell differentiation, but also RAG expression (Table 1.4; [207, 208]). E2A is a member of a large family of transcription factors that share homology in their basic DNA binding domain and their helix-loop-helix dimerization motif (bHLH; [209]). bHLH proteins bind DNA at consensus sequences called E boxes and exist as homo or heterodimers that are ubiquitously expressed (reviewed in [210]). The E2A gene was identified based on the ability of its alternatively spliced gene products, E47, E12, and E2-5, to recognize μE sites in several Ig transcriptional regulatory elements, including HC and LC enhancers, and are thought to have a critical role in Ig transcription (reviewed in [211]). E2A protein complexes have surprising stage- and lineage-specific function in B cell development. In place of this specificity being conferred by expression pattern, it is conferred by two key mechanisms: 1) stage- and lineage-specific binding of these proteins to their cis elements and 2) the ability of family members to homo and hetero-dimerize in certain defined combinations or with stage-specific partner proteins in different lineages and developmental stages.

That the E2A factor is both an important control point in B-cell differentiation and in RAG regulation is suggested by several observations [207, 208]. In particular, E2A−/− mice have a block earlier than RAG1 and RAG2−/− mice (fraction A; see Figure 1.7) and fetal liver pro-B cells of E2A−/− mice completely lack mRNA transcripts for other B-cell specific genes including λ5, CD19, and have reduced expression of mb-1 mRNA [208]. Further suggesting the specific importance of the E2A in B cell development is the observation that E2A−/− mice also lack or have reduced expression of various Ig germline transcripts, consistent with their critical role of E2A in Ig transcriptional regulation [207]. The role of E2A in transcriptional activation of not only the Ig loci but the RAG locus was also suggested in a study by Schlissel et al., in which E47 (an E2A isoform) cDNA clone transfected in a pre-T cell line significantly stimulated expression of not only endogenous germline HC genes, but also
endogenous RAG1/2 while at the same time inducing V(D)J recombination at the Ig HC locus [212]. Further evidence for a role of the E2A products in the transcriptional activation of the RAG locus is the fact that E2A+ mice also completely lack RAG1 transcripts [208]. This also suggests that the RAG1 promoters and/or enhancers have E box DNA-binding motifs for the E2A proteins to confer transcriptional activation of this locus.

Some transcription factors in early B cell differentiation may be B cell-specific RAG repressors, rather than activators (Table 1.3). For example, the DNA binding of E2A may be lineage and stage-specifically regulated through its association with a dominant negative regulator protein Id (reviewed in [210]). The mRNAs for the Id factor are only expressed in pro-B cells (co-expressed with the first RAG expression, predominantly in Hardy fraction A, and at low levels in fraction B cells; Table 1.1), and are absent in more differentiated precursors. Id has a HLH domain but lacks the basic aa region required for DNA-binding domain activity, thus inhibiting other HLH family members by forming non-functional heterodimers. By regulating E2A, Id is therefore an indirect negative regulator of RAG. This conclusion is supported by studies of transgenic animals constitutively expressing Id. These animals have a phenotype that is very similar to the E2A+ mice, including downregulated levels of E2A transcripts and all other E2A target genes, including strikingly reduced levels of RAG1/2 [213]. A zinc-finger protein, ZEB, has also been implicated in inhibiting E2A function [214], and thus, in a fashion similar to Id, may also indirectly repress RAG transcription.
1.2 Regulation of RAG expression: a cellular and molecular perspective

Because some of the genes I identify in my screens may be involved in the regulation of RAG expression, I will discuss in detail the current understanding of RAG regulation both at both the cellular and molecular levels. In general, the regulation of RAG expression is very precise as well as complex. At the cellular level, RAG expression is differentially regulated throughout development by various different combinations of extrinsic and intrinsic factors including cell lineage, differentiation, cell cycle, and various signaling pathways such as antigen receptor signaling, cytokine induction and through costimulatory molecules. At the molecular level, RAG is controlled by transcriptional and post-transcriptional mechanisms. In Tables 1.3 and 1.4, the various extrinsic influences reported to be involved in negatively or positively regulating RAG1 and/or RAG2 expression (either constitutively or inducibly) are listed. Figure 1.9 shows a model of signaling-mediated RAG induction/downregulation at various stages of B lymphocyte development. Below, I will briefly review what is known about extrinsic/influences involved in RAG regulation at different stages of lymphocyte development.

1.2.1 Cellular extrinsic and intrinsic influences regulating RAG expression

1.2.1.1 Regulation of RAG expression “waves 1 and 2” in antigen-independent phases of lymphocyte differentiation

As mentioned earlier, RAGs are induced and downregulated twice in early lymphocyte development (Figure 1.4; [41, 53]). This differential regulation of RAG in progenitor/precursor populations can occur in two ways: via multiple RAG regulatory factors, but also potentially via biphasic regulation by the same factor (Tables 1.3 and 1.4). One factor that may regulate RAG in early lymphocyte development (either directly or indirectly) is IL-7 (see section 1.1.6.5). IL-7 appears to have biphasic effects on RAG expression in lymphocyte development; in particular apparently differential effects between T and B development, in pro-, pre-, and GC mature B cells, and between mouse and human. CD19 may also have a role in conjunction with IL-7 in RAG regulation in early in B cell development (Table 1.4).

Other cytokines may also be involved in inducing/altering RAG expression in early B cell populations, either by working independently from, or in synergy (or redundancy) with IL-7 (Tables 1.3 and 1.4).
### Table 1.3

Extrinsic influences involved in inhibitory regulation (termination or down-regulation) of RAG1 and/or RAG2 expression

<table>
<thead>
<tr>
<th>Category of negative regulator: extrinsic in vivo or in vitro influence</th>
<th>Direct intrinsic target receptor(s)/pathway(s)/gene(s) affected</th>
<th>Possible downstream intermediate effectors and/or other targets besides RAG affected (specificity of effect)</th>
<th>Lineage/stage of differentiation effects on RAG observed/reference</th>
<th>Kinetics of effect on RAG expression</th>
<th>Comment on possible molecular mechanism(s) involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Antigen-receptor signaling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in vitro: soluble (Fab)² anti-μ cross-linking (BCR ligation)</td>
<td>BCR</td>
<td>Specific effect: mRNA expression of various downstream targets assessed not affected including early B-cell specific genes TdT, λ5, VpreB, IL-7R</td>
<td>Mouse immature Eu-Nmyc transformed B cell line (Ma et al., 1992)</td>
<td>mRNA reversible downregulation, &gt;10 fold BY 1H; increases peaking at 6h-Northern</td>
<td>transcriptional, no mRNA stability differences by ActD analysis</td>
</tr>
<tr>
<td>in vitro: soluble or plate-bound anti-CD3 cross-linking (TCR ligation)</td>
<td>TCR</td>
<td>Besides RAG, CD4 and CD8 mRNAs also downregulated; other targets assessed including TCR, MHCI not affected</td>
<td>Mouse DN and DP thymocytes (Turka et al., 1992)</td>
<td>RAG1/2 mRNA termination-Northern</td>
<td>?</td>
</tr>
<tr>
<td>in vitro: plate-bound Mab anti-β TCR cross-linking (TCR ligation)</td>
<td>TCR</td>
<td>PKC pathway mediated; specific effect: mRNA expression of various downstream targets assessed such as CD8, CD4, TCR, MHCI and II not affected</td>
<td>Mouse CD4⁺ CD8⁺ fetal thymocytes (Takahama and Singer, 1992)</td>
<td>3-4 fold RAG1/2 mRNA downregulation-Northern</td>
<td>unchanged transcription rates as determined by run-on analysis post-transcriptional: transcript desabilization; de novo dependent</td>
</tr>
<tr>
<td>in vivo: positive selection</td>
<td>TCR</td>
<td>Downstream target affected: increases in CD69 mRNA expression</td>
<td>anti-MHC class I H-2b TCR bred onto H-2b background, double transgenic mice (Brandle et al., 1992)</td>
<td>only RAG1 mRNA assessed; ~8 fold downregulation in positively-selected (DP TCR⁺) relative to non-selected (DP TCR⁺) thymocytes, no RAG downregulation in thymi of single transgenic mice i.e. bred on non-selecting background</td>
<td>?</td>
</tr>
<tr>
<td>in vivo: positive selection (immunization of 3-83 κ-chain transgenic mice with 3-83 specific high and low-affinity phage-displayed ligands)</td>
<td>BCR</td>
<td></td>
<td>Mouse mature splenic B cells from immunized 3-83 transgenic mice (Hertz et al., 1998)</td>
<td>Termination of RAG1/2 expression and endogenous λ excision products as compared to unimmunized mice (assessed 4-5 days post-immunization): no effect with low-affinity specific phage-displayed ligand</td>
<td>?</td>
</tr>
<tr>
<td>ex vivo: soluble intact anti-κ or anti-λ and (Fab)² anti-κ or anti-λ (BCR ligation)</td>
<td>BCR</td>
<td></td>
<td>Normal mouse mature splenic B cells (Hertz et al., 1998)</td>
<td>RAG1/2 mRNA downregulation</td>
<td>?</td>
</tr>
<tr>
<td>ex vivo: soluble intact anti-κ or anti-λ and (Fab)² anti-κ or anti-λ (BCR ligation)</td>
<td>BCR</td>
<td></td>
<td>Human GC mature B cell centrocytes (Meffre et al., 1998)</td>
<td>inhibition of RAG re-expression by IL-4 and LPS stimulation as assessed by comparing IL-4 + LPS alone with IL-4, LPS, and either anti-κ or anti-lgκ co-incubation</td>
<td>?</td>
</tr>
<tr>
<td>2. Kinases and phosphatases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in vitro: shift from permissive to non-permissive conditions (34-39,5°C)</td>
<td>Inactivation of Abelson Tyrosine Kinase</td>
<td>Increased NF-κB/rel expression and light chain rearrangements</td>
<td>AMuLV-ts (temperature sensitive) mutant pre-B cell lines (Chen et al., 1994)</td>
<td>10-20 fold reversible RAG1/2 mRNA increases peaking at ~4-6h after switch to non-permissive conditions</td>
<td>transcriptional repression: stability unaffected as determined by Act: de novo protein independent as assessed by cycloheximide pre-treatment; cell cycle independent</td>
</tr>
<tr>
<td>in vitro: P34cdc2kinase</td>
<td>Phosphorylation of rag2 aa Threonine 490</td>
<td></td>
<td>Cell-free phosphorylation substrate assay and cell cycle analysis of mouse proB HAFTL cell line and DN thymocytes (Lin et al., 1993, 1994, Li et al., 1996)</td>
<td>RAG2 protein downregulation preferentially in Gy/S; RAG1/2 mRNA and RAG1 protein expression not affected</td>
<td>post-translational: cell-cycle dependent protein desabilization via phosphorylation of Thr490</td>
</tr>
<tr>
<td>in vitro: calyculinA</td>
<td>Inhibition of PP1 and PP2 phosphatases</td>
<td></td>
<td>Mouse pro-B, pre-B and mature B and T cell lines (Casillas et al., 1995)</td>
<td>RAG1/2 mRNA increases transient: peak (&gt;20-fold for RAG1 and 3-5 fold for RAG2) between 4-5 h post-stimulation in mature B cells and between 1-2h post-stimulation in T cells</td>
<td>?</td>
</tr>
<tr>
<td>Category of negative regulator: extrinsic in vivo or in vitro influence</td>
<td>Direct intrinsic target/receptor(s)/pathway(s)/gene(s) affected</td>
<td>Possible downstream intermediate effectors and/or other targets besides RAG affected (specificity of effect)</td>
<td>Lineage/stage of differentiation effects on RAG observed/ reference</td>
<td>Kinetics of effect on RAG expression</td>
<td>Comment on possible molecular mechanism(s) involved</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>3. Cytokines and activators of signaling intermediates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ex vivo: IL-4</td>
<td>IL-4R</td>
<td>Specific for TdT and RAG; TdT in FM GC cells downregulated, other human B cell specific markers such as α, γ, 1, 4, and Igl expression not affected</td>
<td>Human GC mature B cell centrocytes (Nussensweig, unpublished)</td>
<td>RAG1/2 mRNA downregulation-RT-PCR</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ex vivo: recombinant IL-7</td>
<td>IL-7R</td>
<td>Downstream targets affected: TdT expression downregulated, CD19 expression increased; fairly specific effect because expression of other B cell specific genes not affected including IgM, κ, λ, mb-1, and B29</td>
<td>Pre-B cells undergoing LC rearrangements from Eq- hsc2 transgenic mice (Rolink et al., 1993)</td>
<td>RAG1/2 mRNA downregulation-RT-PCR</td>
<td>post-transcriptional</td>
</tr>
<tr>
<td>in vitro: phorbol ester TPA</td>
<td>PKC translocation</td>
<td>?</td>
<td>Mouse pre-T line CEM, pre-B cell line REH, and DN thymocytes assessed (Ncable et al., 1992)</td>
<td>RAG1/2 mRNA termination-northern</td>
<td>transcriptional repression and transcript destabilization; de novo independent</td>
</tr>
<tr>
<td>ex vivo: phorbol ester PMA</td>
<td>PKC translocation</td>
<td>Downstream targets: CD4 and CD8 mRNA expression specifically terminated</td>
<td>Mouse CD4⁺ CD8⁺ fetal thymocytes (Takahama and Singer, 1992)</td>
<td>RAG1/2 mRNA termination-northern</td>
<td>no changes in transcription rates as determined by run-on analysis post-transcriptional: transcript destabilization; de novo dependent</td>
</tr>
<tr>
<td>in vitro: phorbol ester PMA or calcium ionophore A23187 alone, and PMA+A23187</td>
<td>PKC translocation and/or extracellular calcium flux</td>
<td>?</td>
<td>Mouse pre-B RAG⁺ cell line 1-8 (Meinetski et al., 1991)</td>
<td>RAG expression not assessed, but 5-8 fold decreases in recombinase activity by extrachromosomal substrate assays</td>
<td>?</td>
</tr>
<tr>
<td>ex vivo: phorbol ester PMA and calcium ionophore ionomycin</td>
<td>PKC translocation and/or extracellular calcium flux</td>
<td>Specific effect: CD8, CD4, TCR, MHC expression not affected</td>
<td>Mouse DN and DP thymocytes (Turka et al., 1992)</td>
<td>RAG1/2 mRNA termination-Northern</td>
<td>?</td>
</tr>
<tr>
<td>4. Transcription factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in vivo: targeted disruption of Id1 gene</td>
<td>competitive inhibition of E2A binding? (dimerization with BHLH)</td>
<td>8-fold decreased mRNA expression of E2A; other downstream targets affected: increased expression of Cq, Ck, Iμ, mo, and λ5 transcripts</td>
<td>ldl transgenic IgM⁺ B220⁺ BM B cells due to developmental block all are pro B cells (Sun, 1994)</td>
<td>25-50 fold decreases in RAG1/2 mRNA-RT-PCR</td>
<td>?</td>
</tr>
<tr>
<td>5. Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in vivo: targeted disruption of RAG2 locus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>post-translational: increased protein half-life assessed by pulse labeling</td>
</tr>
</tbody>
</table>
### Table 1.4
Extrinsic influences involved in the positive regulation (induction/upregulation) of RAG1 and/or RAG2 expression

<table>
<thead>
<tr>
<th>Category of positive regulator: extrinsic</th>
<th>Direct intrinsic target receptor(s)/pathway(s)/gene(s) affected</th>
<th>Possible downstream intermediate effectors and/or other targets besides RAG affected (specificity of effect)</th>
<th>Lineage/Stage of differentiation effects on RAG observed/reference</th>
<th>Kinetics of effect on RAG expression</th>
<th>Comment on possible molecular mechanism(s) involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Antigen-receptor signaling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in vivo: soluble Fab(α)ß anti-µ cross-linking (BCR ligation)</td>
<td>BCR</td>
<td>PKC pathway early, intermediate, and late mature B signaling events appear normal; multiple secondary rearrangement events including IgM intermediate require upregulation for correction of non-productive rearrangement instead of receptor editing autologous specificity?</td>
<td>Human mature B cell line OCI LYP-C3P (Verkoczy et al., 1995)</td>
<td>4-6 fold reversible increases in RAG1/2 mRNAs assessed by RT-PCR (RAG1 mRNA not assessed)</td>
<td>transcriptional and post-transcriptional (6 fold increase as assessed by nuclear run-on; 2 fold in stability by Act-D, de novo synthesis dependent as determined by cyclohexamide pre-treatment)</td>
</tr>
<tr>
<td>ex vivo: anti-k or anti-idiotypic MAb crosslinking (BCR ligation)</td>
<td>BCR</td>
<td>Immature B cells from transgenic and normal mice (Hertz et al., 1997)</td>
<td>?</td>
<td>3-5 fold increases in RAG2 mRNAs assessed by RT-PCR (RAG1 mRNA not assessed)</td>
<td>?</td>
</tr>
<tr>
<td>in vivo: negative selection; autoantigen encounter (receptor editing)</td>
<td>BCR with autoantigen specificity</td>
<td>Double transgenic &quot;centrally deleting&quot; bone marrow immature B cells; HC+LC, 3-83 idiotypic-specific, anti-H2K(α)β transgensics bred onto either H2K(α)β (high affinity), H2K(β) (low affinity), or H-2K4 (non-autoantigenic, non-deleting) transgenic backgrounds (Tlega et al., 1993)</td>
<td>?</td>
<td>3-5 and 8-11 fold increases in RAG1/2 mRNAs (as assessed by RT-PCR) in 3-83 idiotypic FACS sorted BM cells from high-affinity centrally-deleting mice relative to low affinity and non-deleting mice, respectively</td>
<td>?</td>
</tr>
<tr>
<td>in vivo: negative selection (2°CARE-α rearrangements)</td>
<td>Autoantigen specific for transgenic TCR</td>
<td>αβ TCR transgenic mice (Brandle et al., 1992)</td>
<td>?</td>
<td>2 orders of magnitude increases in RAG1/2mRNA in high versus low-expressing transgenics-RTPCR</td>
<td>?</td>
</tr>
<tr>
<td>2. Activators of signaling intermediates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in vivo: theophylline or caffeine</td>
<td>Inhibition of cAMP phosphodiesterase cAMP (resulting in cAMP accumulation)</td>
<td>intermediate: PKA pathway</td>
<td>1-8 mouse preB cell line (Menetski et al., 1991); other early lymphoid cell lines assessed to have increased recombine activity by extrachromosomal substrate assay</td>
<td>reversible RAG1 mRNA upregulation:10 fold increases, peak at 6h, assessed by Northern</td>
<td>transcriptional only</td>
</tr>
<tr>
<td>in vivo: theophylline</td>
<td>Inhibition of cAMP phosphodiesterase cAMP (resulting in cAMP accumulation)</td>
<td>intermediate: PKA pathway</td>
<td>Mouse pre-B cell line REH and pre-T cell line CEM (Neale et al., 1992)</td>
<td>RAG1/2 mRNA induction/upregulation:2 fold increase, peak at 1h, assessed by Northern</td>
<td>transcriptional and post-transcriptional, de novo synthesis independent</td>
</tr>
<tr>
<td>in vivo: BR-cAMP (cAMP analogue)</td>
<td>Directly raises intracellular cAMP levels</td>
<td>intermediate: PKA pathway</td>
<td>1-8 mouse pre-B cell line assessed (Menetski et al., 1991)</td>
<td>4-5 fold increases in recombine activity as assessed in extrachromosomal substrate assay</td>
<td>?</td>
</tr>
<tr>
<td>in vivo: forskolin</td>
<td>Activation of adenylate cyclase</td>
<td>intermediate: PKA pathway</td>
<td>1-8 mouse pre-B cell line assessed (Menetski et al., 1991)</td>
<td>15-25 fold increases in recombine activity as assessed in extrachromosomal substrate assay</td>
<td>transcriptional and post-transcriptional, de novo synthesis dependent</td>
</tr>
<tr>
<td>in vivo: phorbol ester (PMA or TPA)+ calcium ionophore (ionomycin)</td>
<td>PKC translocation and extracellular calcium flux</td>
<td>intermediate: PKA pathway</td>
<td>Human mature B cell line OCI LYP-C3P (Verkoczy et al., 1995)</td>
<td>6-8 fold reversible increases in RAG1/2 mRNA upregulation peaking at 6h</td>
<td>?</td>
</tr>
<tr>
<td>3. Cytokines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in vivo: IL-7</td>
<td>IL-7R ligation</td>
<td>intermediate: PKA pathway</td>
<td>Mouse thymocytes from fetal thymic organ cultures (Muegge et al., 1993)</td>
<td>RAG1/2 mRNA induction</td>
<td>?</td>
</tr>
<tr>
<td>in vivo: IL-2</td>
<td>IL-2R ligation</td>
<td>Pre-B cells undergoing HC rearrangements from Eβ1c2 transgenic mice (Rolink et al., 1993)</td>
<td>Transient upregulation in RAG1/2 mRNAs</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

51
<table>
<thead>
<tr>
<th>Category of positive molecular effect</th>
<th>Direct intrinsic regulator/extrinsic target or pathway(s) involved in vivo or in vitro</th>
<th>Lineage/stage of cell involved</th>
<th>Possible downstream effect</th>
<th>Lineage/age of other target genes or pathways involved</th>
<th>Kinetics of effect on RAG expression</th>
<th>Comment on possible molecular mechanism involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cytokines, IL-3, IL-4, IL-6, IL-7</td>
<td>human pro-B cell line F1.g3.44</td>
<td>RAG-1 mRNA induction detectable within 12h and peak at 24h</td>
<td>RAG-2 mRNA induction</td>
<td>Absence of transcriptional activation</td>
<td>RAG-1 mRNA induction</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. IL-4, IL-6, IL-7</td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. IL-4, IL-6, IL-7</td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Transcription factors</td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Other</td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mouse GC, human B cells (Takada et al., 1997)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The table contains a mix of text and numbers, indicating various molecular effects and mechanisms. The content appears to be a detailed analysis of molecular interactions, possibly related to immunology or cell biology, with specific references to experimental conditions and outcomes.
Figure 13. Hypothetical model of signaling-mediated regulation of the four RAG expression "waves" throughout B lymphocyte development. In brief, RAG expression may first be induced in pro-B cells either by ligation of the surrogate light chain (SL) "early protein" (p130 or p55) complex with signals delivered from stromal cells or and/or by ligation of the IL-7 cytokine receptor . RAG expression may then be sustained in pre-B cells until rearrangements at the IgH have been completed after which stromal cell signal-SL/early protein complex interactions may then downregulate RAGs until a signal for light chain rearrangement is received. This signal may be ligation of the pre-B cell receptor which then elicits a second wave of expression. Immature bone marrow (BM) BCR+ B cells likely still express low levels of RAG. Depending on differences in ligand affinity, reflecting the autoreactivity of the particular clone, RAG may be differentially regulated in response to cross-linking antigen . Immature B cells in the bone marrow may undergo a third wave of RAG expression in order to undergo receptor editing. Cells which have edited their receptors are then positively selected by another round of antigen encounter . In cells that have further differentiated into mature germinal centre (GC) B cells, a fourth wave of RAG expression occurs, presumably post-hypermutation, in centrocytes of the light zone. The fourth wave of re-expression appears to be helper T-cell and/or FDC-mediated. Signals including CD40/CD40L co-stimulatory interactions (either with FDC or helper T cells) and the cytokines IL-4 and/or IL-7 (see section 1.2.1.3 for details) may be required. Subsequently, high affinity BCR-Ag interactions positively select centrocytes and shut off RAG expression. This is unlike the receptor editing situation in BM where presumably higher affinity BCR-Ag interactions are autoreactive and stimulate RAG expression [226]. Alternatively, low-affinity BCR-Ag interactions may allow RAG expression to be sustained or perhaps even upregulated, thereby allowing revision of receptors via secondary rearrangements for the purpose of another round of FDC selection. After reencounter with antigen and positive selection , RAG mRNAs are shut off and centrocytes differentiate into plasma or memory B cells. Not every cell undergoes four potential waves of expression. For example, a single cell may actually undergo two waves (if selected immediately in both the BM and GC without rearrangement), three waves (if selected immediately in one of the BM or GC), or four waves (if undergoes secondary rearrangements in the BM and GC). Curved arrows represent rapidly cycling cells. The B lymphocyte development and V(D)J rearrangement classifications are based on the nomenclatures of Melchers/Rolink (reference [12]) and Hardy (reference [22]).
Bone marrow

<table>
<thead>
<tr>
<th>Differentiation Status</th>
<th>Rearrangement Status</th>
<th>Locus Accessibility</th>
</tr>
</thead>
<tbody>
<tr>
<td>lymphoid stem cell</td>
<td>H: G</td>
<td>H: closed</td>
</tr>
<tr>
<td></td>
<td>L: G</td>
<td>L: closed</td>
</tr>
</tbody>
</table>

pre-B (Melchers/Robink)
~ Fraction A (Hardy)

- H: D-J
- L: G
- H: open
- L: closed

First "wave": RAG induction

- VpreB1
- VpreB2
- early progenitor

RAG↓

- VpreB1
- VpreB2
- μ, λ

Second "wave": RAG↓

RAG↓

Germinal Centre

- "positive selection" (non-autoactive BCR-Ag interaction)
- RAG shut-off

- "negative selection" (autoactive BCR-Ag interaction)

- Third "wave":
- RAG↑

- secondary light chain rearrangements
- ("receptor editing")

- "positive selection" (non-autoactive sig-Ag interaction)
- RAG shut-off

Key:
- upregulation/induction of RAG mRNAs= ■
- down-regulation of RAG mRNAs= □
- biphasic response of RAG mRNAs= △
Cytokines important for in vitro RAG induction have been assessed in detail in a human pro-B cell line [215]. RAG expression and concomitant recombinase activity in this line are activated by co-culture of the bone marrow-derived stromal cell line PA6 in the presence of cytokines. Among the cytokines, IL-3, IL-6, and IL-7, but not IL-2, IL-4, SCF, or GM-CSF was found to induce RAG expression. In this system, IL-3, IL-6, and IL-7 were found to exert their effects synergistically. Receptors on stromal cells may be responsible for interactions with pro-B surface molecules in RAG induction [216]. One such candidate molecule has been cloned by a differential display strategy [217] and will be further detailed in section 1.3.1.

One additional possibility is that differential alterations in RAG regulation result from differential signaling through the pre-B cell μ/SLC receptor complexes or in analogous pre-TCR complexes (comprised of the pre-TCα (gp33) TM polypeptide and CD3 components either alone or disulfide-linked to the β chain), both which have been demonstrated to be competent physiological signal transducers [88, 93]. Currently, there is no direct evidence for an effect of ligation of SLC complexes on RAG1/2 expression. However, indirect evidence for pre-antigen receptor signaling in modulating RAG expression comes from observations that manipulation of the PKC pathway in pre-B, pre-T cell lines, or DN thymocytes i.e. treatment with activators of second messengers of antigen receptor cascades such as phorbol ester 12-O-tetradecanoylphorbol-13-acetate (PMA; a potent translocater and activator of PKC) in combination with the calcium ionophore ionomycin (which allows extracellular Ca²⁺ fluxes), downregulates RAG mRNAs [40, 218]. Conversely, cyclic AMP (cAMP)-inducing agents including forskolin, caffeine, and theophylline, can upregulate RAG1/2 expression in pre-B cells with a concomitant increase in V(D)J recombinase activity [219]. Therefore RAG up- and downregulation in precursor lymphocytes may be mediated through separate pathways: upregulation through a cAMP-dependent protein kinase A (PKA) signaling pathway, and downregulation through the PKC pathway. Downregulation of RAG may also include a pathway in which the protein tyrosine kinase v-abl is involved, since temperature-dependent inactivation of v-abl is followed by RAG upregulation, NF-kB/rel expression, and LC rearrangements in temperature-sensitive AMulv transformed cells [220]. Because this is not a normal cellular protein, it is hard to assess the physiological significance of this latter effect i.e. if c-abl can negatively regulate RAGs in the same fashion.
Pre-antigen receptor-independent combinations of extrinsic and intrinsic factors may also be involved. For example, one interesting observation is that incubation of the RAG-downregulated intermediate CD25+ thymocyte population spontaneously upregulates RAG expression ten fold in medium (relative to direct ex vivo samples), suggesting RAG down-regulation in vivo is mediated by a reversible negative signal that was not present in the medium [41]. In contrast, the same effect could not be observed in the earlier, high RAG-expressing CD25+ population, suggesting differential regulation of RAG expression in these two subpopulations. A similar upregulatory effect was also found upon incubation of day 19 CD4+ fetal thymocytes in medium alone [221].

1.2.1.2 Antigen receptor-mediated differential regulation of RAG expression “wave 3” in primary lymphoid populations

The continued expression of RAG in antigen receptor expressing lymphocytes suggested that antigen receptor expression alone was not sufficient to shut off their expression, thereby leading to studies of signaling-mediated regulation of RAG through antigen receptors [40, 62, 222, 223]. Surprisingly, two very different outcomes can result under different situations. On one hand, numerous reports have provided strong evidence for the decrease or termination of RAG expression upon signals generated through intact antigen receptors. For example, in vitro, Ma et al. found that cross-linking the BCR of sIg+RAG+ B cell lines from Eμ-N-myc transgenic mice with polyclonal, soluble F(ab)'2 anti-μ reversibly and specifically down-regulated RAG1/2 mRNAs. Similar observations have been made by Turka et al. and Takahama and Singer with signals delivered through thymocyte TCRs [40, 221]. Furthermore, stimulation with agents that activate the phorbol ester and ionomycin also led to significant decreases in RAG mRNAs in the above reports. Several reports in transgenic systems have strengthened the observation of Turka et al. by providing in vivo evidence that specific engagement of TCR cortical thymocytes with self-MHC during positive selection results in the downregulation of RAG expression [222, 224-226].

On the other hand, a third "wave" of RAG expression can be induced or upregulated via antigen receptor signaling. In a pivotal study, Tiegs et al. demonstrated that in autoimmune double-transgenic (centrally-deleting) mice, an autoreactive BCR+ bone marrow B cell sub-population already expressing
RAG mRNAs respond to autoantigen by further inducing their RAG mRNA levels [227]. This observation was coupled with two other key observations. First, centrally deleting mice, but not non-deleting controls, had circular DNA excision products (byproducts of deletional rearrangement events; see Figure 1.10) in bone marrow, suggesting an increased amount of secondary rearrangement events. Secondly, IgM+ B cells that escaped deletion in the centrally deleting mice expressed \( \lambda \) chains with significantly higher frequency than those in non-deleting mice, suggesting that encounter of self-reactive B cells (bearing 3-83 anti-H-2\(^k\) specific \( \kappa \) transgenes) with H-2\(^k\) autoantigen generate endogenous \( \lambda \) chain rearrangements in bone marrow. These observations prompted Tiegs et al. to suggest the "receptor editing" model ([227]; reviewed in [228, 229]). In this model, non-autoreactive clones have their RAG mRNAs shut off and no further rearrangements occur. However, in situations where a B cell has undergone "autoreactive" primary rearrangements, RAG1/2 upregulation allows further HC or LC secondary rearrangements. Such rearrangements in autoreactive B cells can then generate new receptor specificities, thus replacing or "editing" a productively rearranged, but autoreactive HC or LC. By re-initiating endogenous V(D)J rearrangements and displacing autoreactive transgenic receptors, B cells can be spared from central/peripheral clonal anergy and/or clonal deletion, the two established tolerance mechanisms known to operate in mice with transgenic autoantibody genes [227]. A similar receptor editing mechanism has been reported by Weigert and co-workers in mice which spontaneously produce high levels of dsDNA antigens and are engineered to carry \( \alpha\)-dsDNA antibody transgenes [230-233]. In hybridomas from these MRL/lpr transgenic mice, not only are the LC transgenic \( \alpha\)-dsDNA specificities replaced by non-autoreactive endogenous \( \kappa \) and \( \lambda \) LC genes, but also from elimination of the autoreactive transgenic HC gene with endogenous HCs. This HC receptor editing is believed to occur via intrachromosomal recombination using cryptic RSS and is termed "V gene replacement". In this system, high-affinity \( \alpha\)-dsDNA transgenes promote deletion or receptor editing whereas lower affinity transgenes promote anergy. Finally, recent studies carried out by Hertz and Nemazee have also demonstrated BCR ligation-induced receptor editing in both ex vivo IgM+IgD+ transgenic and in non-transgenic bone marrow B cells, as measured by upregulation of RAG expression and LC gene
Figure 1.10. Three possible mechanisms of RAG-mediated secondary rearrangements for the purpose of "receptor editing".

In all examples, numbers are arbitrarily designated. Dashed lines indicate rearrangement events. White, gray, and black triangles represent standard 23 spacer RSS, standard 12 spacer RSS, and RSS containing an embedded heptamer, respectively. A. V gene replacement at the IgH locus using the heptamer from a cryptic 12 RSS embedded at the 3' end of the original V(D)J rearrangement. This mechanism can occur on the human or mouse heavy chain Ig locus. The numbers in brackets indicate order of rearrangement events. B. Inactivation of productive κ locus primary rearrangement using the kappa deleting element RS. This silencing can occur by two possible types of RS recombination: Vκ-RS (left) or VκJκ-intron-RS (right). The excision RSS products of primary and secondary rearrangement events (where applicable) are not shown to simplify diagrams. Symbols include: (IRS1), intronic recombination sequence 1; (iE) and (3'E), intronic and 3' kappa enhancers, respectively; and (RS), recombining sequence element (aka. kappa-deleting element). C. Replacement of deletional or inversional primary rearrangements by standard deletions or inversional light chain secondary rearrangement events. Shown are four possible rearrangement products of primary and secondary rearrangements at the same mouse kappa locus (depending on whether inversional or deletional rearrangements have occurred). Note that further rearrangements are possible, either on the same allele, or a previously excluded allele, such that multiple rounds of editing are possible. The same phenomenon can be seen at the human lambda locus. Only the region of the locus containing V and J gene segments is shown.
replacements (κ to λ editing) [234]. Many other reports of ongoing in vitro secondary κ and λ rearrangements (both in cell lines and primary cultures) have also been documented [235-240]. Examples of types of secondary rearrangements are shown in Figure 1.10.

Presently, the physiological relevance of receptor editing relative to deletion and anergy is not known. Statistical models argue that the occurrence of autoreactive sIg+ B cells being "edited" can vary as a function of B cell deletion [228]. Furthermore, a recent study by Retter and Nemazee has argued that in normal, non-transgenic B cell development, receptor editing takes place at a high rate in bone marrow [241]. A second unresolved issue with respect to receptor editing is what the cellular and molecular triggers are in sIg+ populations that result in the differential regulation of RAG expression. Another way of asking this is what constitutes an "autoreactive" specificity versus a "non-autoreactive" specificity and when does the cell decide to "edit" receptors rather than deleting or anergizing B cells. With respect to cellular aspects that influence these outcomes, some possibilities include parameters that are thought to influence tolerance outcomes in other systems (reviewed in [114]. These may include ligand affinity/avidity, presence or absence of co-stimulation, or differences intrinsic to the developmental stage (such as BCR structural differences or differential expression/association of signal transduction intermediates and transcription factors) at which the "editing" events actually occur. In this context, a recent report by Nemazee and co-workers suggests that ligand affinity determines whether the tolerance outcome is receptor editing or anergy/deletion [242].

1.2.1.3 Signaling-mediated regulation of RAG expression "wave 4" in mature GC B cells

While it appears that antigen receptor cross-linking alone in primary lymphoid tissues is sufficient to induce a third wave of RAG expression, other extrinsic signals appear to be required to induce the 4th wave of RAG expression in mature B cells. These include co-stimulatory T-B interactions and cytokines that result from further interaction with helper T (Th) cells or FDCs in the GC microenvironment (Table 1.3). In vitro, RAG1/2 transcripts have been found "re-expressed" in mouse mature GC B cells after culture with various agents in combination with IL-4 including LPS and α-CD40 (CD40/CD40L interactions) [72, 243]. As discussed earlier, IL-7 can also act as an in vitro and
in vivo cofactor in inducing the 4\textsuperscript{th} wave of mouse RAG expression [199]. Because the impairment in RAG induction seen when blocking IL-7/IL-7R interactions is not seen in immunized IL-4\textsuperscript{+} mice, and co-incubation of IL-7 and IL-4 do not additively or synergistically enhance RAG expression, this suggests an IL-4-independent role for IL-7/IL-7R in GCs RAG re-expression [199]. The observation of combined cytokine and co-stimulation-mediated RAG induction in GC B cells is furthered by recent findings that in vitro incubation with T\textsubscript{H}2 clones along with specific antigen can substitute effectively to induce RAG expression [199]. In vivo, induction of RAG1/2 transcripts has been detected in lymph node and spleen GC of immunized mice [71, 72]. Kelsoe's group further characterized the predominant RAG-inducible GC B cell subset to be centrocytes, as defined by positive B7-2 expression [71]. The picture is different in humans where, in direct contrast to studies in mouse splenic GC B cells, Nussenzweig and colleagues have shown that certain cytokines can downregulate RAG mRNAs in human mature B GC cells [82]. For example, in human centrocytes, treatment with IL-4 alone downregulates RAG mRNAs. Furthermore, co-culture of CD40L-expressing human centroblasts with IL-2, IL-4, or IL-10 was found to have no effect on RAG expression. These results suggest potentially different requirements for induction/and or downregulation of RAGs in human and mouse GC mature B cells. Alternatively, because the centrocyte subpopulation assessed in this study already expresses high levels of RAGs (comparable to BM fractions), induction in this population may not be possible and/or may have already undergone secondary rearrangement/selection events. It is also possible that RAG expression in this particular subpopulation may be differentially regulated.

While the combination of cytokine and co-stimulatory signals can upregulate/induce a 4\textsuperscript{th} wave of RAG expression in GC B cells, antigen receptor-signaling appears to have the opposite effect in mouse and human. In a study with human GC cells, ex vivo cross-linking of primary human centrocyte subsets with anti-Ig downregulates RAG expression [82]. In another study by Hertz and Nemazee, ex-vivo IL-4+LPS stimulation of mature splenic mouse B cells re-expresses the RAGs, whereas co-incubation with an anti-Ig inhibits this induction. Based on these findings, the authors suggest that unlike BCR ligation in the bone marrow, BCR ligation in GCs only indirectly influences RAG induction via interaction of antigen-stimulated centrocytes with T\textsubscript{H} or FDC signals and T\textsubscript{H}2-secreted cytokines. In this same study, immunization with high-affinity, but not low-affinity specific ligand, shuts off RAG
expression and secondary λ rearrangements in the 3-83 transgenic system. This therefore argues that BCR signaling-mediated termination of RAG expression may be an affinity-dependent effect, one that is opposite to the affinity requirements of those in bone marrow (where high affinity has a stronger effect on RAG induction than does low affinity ligation [227]). However, I have previously provided direct in vitro evidence that BCR signaling alone (either by BCR ligation or by PMA+ionomycin stimulation) in the human mature GC-derived mature B cell line OCI-LY8 results in increased RAG expression, rather than unchanged or down-regulated RAG levels [171]. One possibility for the discrepant results between studies could be that soluble, lower affinity BCR ligand used in my studies results in upregulated rather than unchanged levels of RAGs, but intact or plate-coated anti-Ig i.e. high affinity BCR ligand, may in fact downregulate RAGs. A caveat with these findings is that their physiological relevance cannot be determined since the upregulation of RAGs could be due to a more general deregulation specific to large cell lymphoma cell lines. Further experiments with different affinity BCR ligands may help to discern the differential regulation of RAG in GCs and “editing” versus “selection” events.

1.2.1.4. A role for RAGs in GC receptor selection/diversification rather than receptor editing?

What role do re-expressed RAG products have in GCs? Two initial candidates were the well-documented mechanisms of class switching and hypermutation, both of which occur in GCs. With respect to hypermutation, based on where the RAGs are expressed (post-hypermutation centrocyte subsets), it is unlikely that they are involved in this process. Further recent evidence against a role for RAG in hypermutation is the extensive mutation of LC transgenes in Ig transgenic RAG1+/mice reconstituted with specific T_H cells and antigen [71, 244]. A potential role for RAG in class switch seemed more likely because of the co-occurrence of class switch and RAG expression specifically in GC light zones upon induction with LPS+IL-4. However, a role for RAGs in class switch is implausible in light of the fact that RAG1 or RAG2+/mice in which a rearranged HC and λ LC has been "knocked-in" at the endogenous unrearranged V(D)J locus are still fully competent to undergo IgG isotype switch recombination in response to LPS+IL-4 stimulation [245, 246].
Recent studies using ligation-mediated PCR (LM-PCR), an assay which detects de novo, RAG-specific dsDNA breaks, have demonstrated the functionality of RAG proteins. For example, our lab has reported that RAG1/2 are re-expressed (as assessed by RT-PCR) and functional (as assessed by LM-PCR), in human mature GC B cells from activated tonsils [247]. Since functional recombinase activity is detectable in GCs [248-250], coupled with the finding that RAG expression is expressed in the light zone (post-hypermutation) GC compartment, one possible role for a fourth RAG expression wave may be in editing receptors, analogous to the third RAG expression wave in the bone marrow. Why would additional editing be necessary in peripheral lymphoid tissues? The most obvious reason is because the randomness of the somatic hypermutation process in HC and LC variable region generate clones expressing autoreactive antigen receptors just as where primary rearrangements in pre-B cells generate autoreactive or non-productive rearrangements in the bone marrow. Evidence for the generation of hypermutation-mediated autoreactive specificities comes from recent studies in which cross-reactive anti-dsDNA autoantibodies were generated in the GCs of normal mice during immune responses to foreign antigen [251]. Another possibility for increased GC RAG expression (suggested by Kelsoe) is as a byproduct of general deregulation in apoptosing centrocytes, since many other pre-B-specific genes, including TdT, VpreB, and λ5 can also be re-expressed in centrocytes [82]. A final, possible reason for RAG expression in GC centrocytes is to enable B cell clones originally with low-affinity receptors to acquire higher affinity receptors after positive selection by FDCs. In this context, GC RAG expression would not be a tolerance-driven "editing" mechanism, but rather a mechanism of further diversifying specificities and creating higher affinity receptors. On this view, GC RAG expression may actively create new autoimmune specificities. The recent studies by Hertz et al. in which high affinity but not low affinity BCR ligation downregulates, rather than upregulates GC B cell RAG mRNAs are consistent with this latter possibility, which has been called "receptor selection" rather than "receptor editing" ([82, 253]; reviewed in [252]).
1.2.1.5 Factors which influence RAG expression throughout lymphocyte development

Some factors may be responsible for controlling RAG expression throughout lymphocyte development. For example, Casillas et al. described data implicating the serine/threonine phosphatases in regulating expression of the RAGs [254]. Using calyculin A, a potent inhibitor of the types 1 and 2A serine/threonine protein phosphatases (PP1 and PP2A, respectively), these investigators found that expression of RAG1/2 mRNA was upregulated in pre-B as well as B and T cell lines. This occurred with a concomitant increase in recombination of a transfected V(D)J joining substrate in all lines tested. These results suggest that such phosphatases are somehow involved in signaling RAG induction throughout lymphocyte development. Both positive regulation from phosphorylation of RAG transcription factors and negative regulation by phosphatases may therefore influence RAG transcription.

1.2.2 The three molecular levels of RAG regulation

At present, there is still relatively little known about the molecular control of RAG1/2 expression. There is evidence, however, that expression of these genes is regulated at three distinct levels: transcriptional, post-transcriptional, and post-translational. The latter level includes a phosphorylation-dependent cell cycle mechanism of controlling RAG2 protein expression, a mechanism shown not to affect RAG1 or RAG2 mRNA expression [63, 64, 255]. Furthermore, the post-translational control of RAG proteins is not related to genes I would have any likelihood of isolating by DD RT-PCR. For these reasons, this level will not be discussed further here; the reader is consulted to relevant reviews on the matter [55]. This section will focus on reviewing the various reports which have contributed to a preliminary molecular understanding of variables involved in regulating RAG1/2 mRNAs. However, in order discuss molecular aspects of RAG mRNA regulation, the general identification, genomic structural organization, and conservation of the RAG locus will first be reviewed.

1.2.2.1 Genomic organization and conservation of the RAG locus

The RAGs are located on human and mouse chromosomes 11 and 2, respectively. The human RAG1/2 locus was localized to the short arm of chromosome 11 (11p12-13) by fluorescence in situ
hybridization (FISH) analysis [256-258]. This region has been the subject of intensive study due to the presence of numerous disease loci in this region. Based on a recent high-resolution map (within a one megabase region) of 11p12-13 using human-hamster somatic cell hybrid PCR analysis, the human RAG locus maps between CD44 (Pgp-1) and D11S9. The RAG locus has been characterized in a number of species, including human [256], rabbit [70], mouse [44], chicken [48], shark [259], and frog [260]. In all examples, RAG1/2 has a conserved structural genomic organization. Although RAG1 and RAG2 have no sequence homology to each other and thus have not arisen via gene duplication, they are closely juxtaposed in a tail to tail configuration and are convergently transcribed [261]. The entire coding regions of both RAGs are contained within a single, large exon upstream of one or more small exons containing portions of the 5' untranslated (UTR) region.

Recent work from our laboratory represents the most detailed characterization of the genomic organization of any RAG1 and RAG2 locus, including the transcriptional start site (TSS) and promoter regions [262]. Figure 1.11 summarizes the genomic organization of the human RAG locus. Briefly, the human RAG1 locus consists of two exons (the coding exon is ~6.5 kb) separated by an intron of 5.2kb. Due to the absence of an upstream TATA box and an Initiator element, transcription initiates from a cluster of start sites. The RAG1 gene generates a 6.6 kb transcript encoding a 119 kDa protein (1041 amino acids). The RAG2 locus consists of at least three exons. Exon 1A, located 3.9 kb upstream of the coding exon (exon 2) contains the major TSS. Exon 1B, located 0.7 kb upstream of exon 2, contains the minor start site. Exons 1A and 1B are alternatively spliced exons whose usage has been observed in both B and T lymphocytes. Similar to the RAG1 locus, the RAG2 locus lacks a TATA box within the normal distance (-30) upstream of the TSS. An Inr element however, is present at exon 1A. In humans, a 16 kb intergenic region separates RAG1 from RAG2.

While the overall genomic structure of the RAG locus is highly conserved across species, notable structural differences exist in the intervening sequences. First, there is considerable variation in the size of the intergenic region. For example, in mouse, human and frog, RAG1/2 are separated by 7, 16 and 2 kb intergenic regions, respectively. Second, the size of the RAG1 intron is 2.1 kb in frog; 5 kb in mouse and human [260, 263]. This suggests that no significant evolutionary pressures obligated
Figure 1.11. Genomic organization of the human RAG locus. Adapted from Zarrin et al., 1997, reference [262].
maintenance of these regions as transcriptional regulatory control spots unless within the common 2 kb fragment.

1.2.2.2 Inter- and intra-species RAG2 message diversity

Genomic probes specific for human RAG1 and RAG2 detect predominant mRNA species of ~6.7 and ~2.2 kb respectively, encoding predicted proteins of 1043 and 527 aa respectively (~119 and ~59 kDa). 5' RACE assays, has determined that subtle differences (undetectable by Northern blot) exists among mRNA species within human cell lines. The absence of a TATA nucleation site in the RAG1 promoter [262] may precipitate minor variations resulting from inaccurate RNA polymerase II positioning. Across species, the size of the RAG1 transcript is about 6.7 kb in most Northern blots. Only one study yielded an additional, smaller RAG1 transcript of unspecified size in mouse thymocytes [64]. RAG2 transcripts, conversely, show a considerable amount of intra-and interspecies diversity. Although the predominant Northern blot transcript isoform in human is ~2.2 kb, there are also minor transcripts of 1.8 and 3.7 kb (Zarrin and Berinstein, manuscript in preparation). Furthermore, Northern blots probed with mouse RAG2 show a predominant isoform of 2.1 kb (the approximate size of the published mouse RAG2 cDNA) and a 3.2 kb minor band [44]. In rabbit, the predominant hybridizing transcript is 4.4 kb [70]. By Northern blot analysis, larger isoforms of RAG2 mRNAs (up to 5 kb) have also been observed in mouse and human (N. Stiernholm and D. Schatz, personal communication). Numerous factors, might explain the presence of multiple RAG2 transcripts within species, including the presence of multiple 5' upstream UTR RAG2 exons, 3' UTR diversity or even cross-hybridization to a gene with homology to RAG2. In human, 3' RACE analysis has determined that RAG2 message variants result from 3' UTR diversity i.e. different usage of polyadenylation sequences in the 3' UTR (A. Zarrin, personal communication). The significance of structural diversity in terms of post-transcriptional regulation will be discussed below.

1.2.2.3 Transcriptional regulation of the RAGs

1.2.2.3.1 Evidence for RAG transcriptional regulation

Data I generated prior to my doctoral studies [75, 264] as well as from others [62, 218] suggests that
transcriptional regulation plays the predominant role in modulating both constitutive and inducible steady-state RAG1/2 mRNA expression levels. With respect to constitutive differences in the human B cell line OCI LY8-C3P, the RAG1 transcription rate is 6-8 fold higher in variants which constitutively express ~20 fold higher RAG1 mRNA levels [264]. With respect to inducible RAG1 mRNA upregulation via either crosslinking of the sIgM receptor (using a soluble, polyclonal anti-μ F(ab)'2 fragment) or by second messenger stimulation (using ionomycin and PMA concomitantly), RAG1 mRNA synthesis rates are co-ordinately and rapidly increased >6 fold [75]. In another study, Ma et al. showed that decreased steady state levels of RAG1/2 mRNA levels induced by sIgM crosslinking in Eμ-N-myc lymphomas were not associated with reduced transcript half-life. This observation implied that these changes in RNA levels were due largely to co-ordinate decreases in transcription [62]. Similarly, in signaling-mediated modulation of RAG mRNA expression in precursor lymphocytes, decreased RAG1/2 steady-state levels can be accompanied by changes in the rates of RAG1/2 mRNA synthesis [218]. Finally, treatment of REH pre-B cells with theophylline, an agent which elevates both intracellular cAMP and RAG steady state mRNA levels [219], also increases RAG1 transcription rates by two fold [218].

1.2.2.3.2 Structural/functional characterization of RAG core promoters

No data had been published concerning the cis-acting regulatory regions of the mouse or human RAGs until 1996, when several groups made advances in isolating and characterizing the RAG promoter regions. Using DNaseI hypersensitivity and 5' RACE assays, Fuller et al. identified a 186 bp region containing an "enhancer-dependent, TATA-less" promoter of the mouse RAG1 gene, which could cause expression of a reporter gene in pre-B and B cells [265]. This study and that of Kurioka et al. for human RAG1, also revealed multiple transcription start sites (TSS), as identified by 5'RACE, primer extension, and RNAse protection assays [266]. Using deletion constructs, Brown et al. and Kurioka et al. further characterized the mouse and human RAG1 5' promoter regions, respectively [266, 267]. In both situations, they found that the 5' promoter region contained a CCAAT box between -110 and -86 that was indispensable for basal promoter activity. Furthermore, Brown et al. demonstrated that a CCAAT binding factor, NF-Y plays a role in the positive regulation of RAG1
Although the NF-Y transcription factor is not lymphocyte restricted, it has been implicated in the lymphocyte-specific expression of the MHCII and IL-4 genes [268, 269, 270 IRF-2, and]. In both studies, the core promoter region as well as 5' flanking regions (up to ~2kb) were active when transiently transfected in lymphoid and nonlymphoid cell lines. Our lab has confirmed the finding of non-tissue specific regulatory regions immediately 5' flanking the human RAG1 promoter, as a luciferase reporter construct containing -1237 to +89 of the major RAG1 TSS exhibits equally high activity in non-lymphoid and lymphoid cell lines [262].

In addition to RAG1, our laboratory has identified the TSS (using RNase protection, primer extension, and 5' RACE methodologies) and the promoter regions (using transient luciferase reporter assays) of human RAG2 [262]. As with the RAG1 promoter, the human RAG2 promoter is TATA-less and its 5' flanking regions (up to 2.7 kb 5' of exon 1A) are active in both lymphoid and non-lymphoid cell lines. Likely, a regulatory element lying outside the immediate 5' flanking region (either the same one is active in RAG1 expression or a distinct element) is involved in the tissue-specific transcriptional regulation of the RAG2 gene. The RAG2 gene initiates transcription from two distinct regions, each with its own unique TSS. One start site constitutes a major TSS, 3.9 kb upstream of the RAG2 coding region. The other constitutes a minor TSS (0.7 kb upstream), defining two exons, 1A and 1B, which alternatively splice in a stage and lineage-independent manner to exon 2 in the mature RNA transcript. Deletion analysis of the human RAG2 basal promoter has revealed that almost all basal promoter activity resides within a 207 bp fragment (-158 to +49 from the major TSS; Fong et al., manuscript in preparation). Subsequent EMSA analysis of a 24 bp fragment from -154 to -134 has revealed the CCAAT/enhancer binding protein (C/EBF) transcription factors bind within this RAG2 positive regulatory region.

In contrast to studies with the human RAG2 promoter, Lauring and Schlissel have recently found that the mouse RAG2 promoter can be transactivated in a tissue specific fashion through binding of a conserved BSAP site within the RAG2 core promoter [271]. In this study, cell lines transfected with BSAP cDNAs were found to activate RAG2 promoter reporter assay constructs containing the putative BSAP -71 to -45 cis region in B cell lines, but not in non-lymphoid cell lines and to a lesser extent in T
cells. Furthermore, this putative BSAP-containing region was found critical for this promoter activity and was occupied in a B cell-specific manner in EMSA and *in vitro* and *in vivo* DMS footprinting assays. Furthermore, these authors found that the mouse RAG2 promoter can be regulated differentially in the T and B cell lineage by distinct, as yet, unidentified transcription factors.

1.2.2.3.3. **The ongoing search for tissue-specific transcriptional regulatory elements flanking the core RAG promoters**

Several approaches have been undertaken to search for RAG tissue-specific regulatory cis elements. In the first approach, Kitagawa *et al.* and Fuller *et al.* have examined the chromatin structure of a 24 kb genomic fragment adjacent to the human and mouse RAG1 promoters, respectively, in various RAG1-expressing lymphoid cell lines by analyzing DNaseI hypersensitivity sites (DHS) [265, 272]. Three DHS which are strikingly similar in mouse and human regulatory regions were identified; two located ~0.8 kb and 0.1 kb upstream of the RAG1 TSS (the latter overlapping the indispensable CCAAT element), the third 1.7 kb downstream of the first exon. However, using transient reporter assays, none of these DHS had any enhancing effects on RAG1 promoter activity. One explanation for this is differing regulatory effects of these regions *in vivo* and *in vitro*, similar to other models such as the β-globin locus control region (LCR) [273, 274]. A second approach has been undertaken in our lab. In this *in vitro* approach, stable transfectants of P1 phage genomic clones containing RAG1/2 and surrounding sequences have been generated, in the human non-lymphoid cell line HeLa (RAG1/2') and the mouse pre-T cell line NFS-70 (RAG1/2') [275]. RAG2 expression was observed in both lines suggesting that this ~75 kb region does not contain the regulatory elements to restrict its expression in lymphocytes. However, RAG1 expression was only detected in NFS-70 suggesting that a transcriptional silencer within this fragment may turn off expression of RAG1 in non-lymphoid cells [275].

Although the above reports may be consistent with the findings of Lauring and Schlissel (that lineage-specific trans-acting factors confer tissue specificity to RAG promoter regions [271]), two recent reports provide convincing *in vivo* evidence that key tissue-specific cis-acting factors lie outside the RAG promoter regions. In these reports, Alt and Nussenzweig’s groups have measured the ability of
cis-regulatory elements to reconstitute RAG1/2 tissue-specific activity in vivo using two distinct assay [276, 277]. Alt's group has assessed complementation of the developmental block in RAG2−/− mice as a readout [276]. In these rescue experiments, several genomic constructs (containing sequences 5' of the RAG2 TSS) that are competent for rescuing RAG2 expression in pre-B cells have been subsequently transfected into RAG2−/− ES cells and tested for rescue of lymphocyte development in RAG2 mouse chimeras [276]. It was found that relative to constructs containing larger portions 5' of the RAG2 promoter, more proximal elements (2-7 kb 5') rescue B, but not T cell development, suggesting the presence of proximal B, and more distal T cell-specific cis regulatory elements. The second methodology, adopted by Nussenzweig's group, involved the generation transgenic reporter mice in which numerous large (up to ~200 kb) Bacterial Artificial Chromosome (BAC) transgenes spanning regions flanking both sides of the RAG locus and containing Green Fluorescent Protein (GFP) and Yellow Fluorescent Protein (YFP) indicators in place of RAG1 and RAG2, respectively [277]. Using this approach, which allows expression of the transgenic constructs to be precisely followed in developing mouse lymphocyte populations by FACS analysis, it was found that expression of RAG1 and RAG2 in T and B cells is regulated by distinct cis elements, all on the 5' side of RAG2. A similar transgenic approach using a BAC/GFP transgenic reporter system has also been used in zebrafish to find a transcriptional RAG1 silencer element between 8 and 12 kb 5' of the RAG1 promoter in Zebrafish [278]. Overall, all reports cumulatively suggest a complex picture of coordinate, tissue-specific RAG transcriptional regulation, in which control is conferred not only by separate T and B transcriptional regulatory factors and/or signaling pathways acting within the RAG promoters, but also by multiple lineage and stage cis-acting regulatory regions outside of the RAG promoters. The results of the above studies will be further addressed in Chapter 5 (General Discussion) in the context of T versus B cell-specific RAG regulation.

1.2.2.4 Post-transcriptional regulation of RAG mRNAs

The stability of intermediate and highly-expressed mammalian mRNA species can vary from half-lives (t₁/₂) of 20 minutes to >24 hours [279]. Tissues and cell lines with high RAG1/2 mRNAs expression (>1000 copies/cell) appear to have RAG t₁/₂ at the unstable end of this spectrum (~30 minutes-2 hours
for RAG1 as assessed in thymocytes, pre-T lines and B cell lines and ~2-4 hours for RAG2 as assessed in the pre-T line CEM) [62, 218]. The fact that RAG2 mRNA has a slightly higher stability than RAG1 mRNA suggests that reduced RAG2 total RNA levels relative to RAG1 are due largely to "tighter" RAG2 transcriptional repressor mechanisms. Consistent with the relative instability of RAG1/2 mRNA is the presence of numerous copies of the pentanucleotide AUUUA sequence (known as A+U rich elements or AREs) in mouse and human RAG1/2 mRNA 3' UTR regions (for example, RAG1 has 13 ARE repeats in its 3.4 kb 3' UTR [4] and A. Zarrin, unpublished data). These sequences are believed instrumental in RNA destabilization mechanisms in mRNAs with rapid turnover rates, such as those of certain proto-oncogenes and cytokines [280]. Because human RAG2 transcript diversity is attributable to 3' UTR regions with varying numbers of AREs (A.Zarrin, unpublished data), it may well be that the t₁/₂ isoform of each is differentially regulated.

The combination of altered gene transcription and message stabilization is generally uncommon [279], but has been observed in a few instances such as the early response genes c-fos [281], c-myc [282] and the IL-2 gene [280]. Studies from our lab and others demonstrates that in addition to alterations in transcription, modulations in transcript stability can account for changes in steady state levels of RAG1/2 mRNA. For example, in both the constitutive and signaling-mediated upregulation of RAG1 mRNA steady state levels in OCI LY8 variants, I have previously also observed involvement of a post-transcriptional component [75, 264]. With respect to constitutive differences in this system, in the high RAG-expressing C3-A11N B cell line (i.e. >1000 copies, cell), the RAG1 t₁/₂ is ~45 minutes, as assessed by measuring RNA accumulation using the transcriptional inhibitor Actinomycin D [264]. This is consistent with ranges observed by others in actively-rearranging RAG-expressing pre-B cell lines (in normal thymocytes and pre-T cell lines, RAG1 seems to be more stable with a t₁/₂ of ~2hr) [62, 218]. However, in the low RAG-expressing OCI LY8 parental variant, which expresses ~20 fold lower RAG1 RNA steady state levels than C3-A11N, the t₁/₂ is 2-3 times shorter (~15 minutes).

With respect to signaling-mediated alterations in RAG1 mRNA steady state levels, I have previously observed that the t₁/₂ of RAG1 RNA increases by ~two-fold (from 45 minutes to 1.5h) in the OCI LY8-C3P mature B lymphoma cell line in response to sIg cross-linking or ionomycin+PMA.
stimulation [75]. Transcript destabilization also plays a role in signaling-mediated downregulation of RAG RNA in early lymphocyte development. For example, in the pre-T cell line CEM, in DN thymocytes and in the pro-B cell line REH, TPA treatment decreases the RAG1/2 mRNA t_{1/2} from two to one hour [218]. Furthermore, a study by Takahama and Singer also argues for a destabilization mechanism in RAG mRNA downregulation as seen in PKC-mediated effects in normal DN and DP thymocytes [221].

In the context of these studies coupled with RAG 3' UTR sequence, it is tempting to speculate that PKC signaling-mediated destabilization is mediated by the de novo synthesis and/or or post-translational activation of ARE-binding proteins. The finding that inducible RAG1 mRNA expression is partly controlled by message stabilization conflicts with that of Chen et al., who reported no differences in RAG1/2 t_{1/2} between non-permissive (high RAG-expressing) and permissive (low RAG-expressing) conditions [220] and by Ma et al. in Eu-N-myc transformed cell lines [62]. The differing results in these independent experimental systems could be attributed to the presence of different regulatory mechanisms, but could also be due to differences in interpretation. For example, Chen et al. also observe a 2-fold increase in stability for RAG2 (and no differences for RAG1) but do not consider this difference significant. However, in a physiological context, with respect to repeated cell divisions, this may cumulatively translate to a substantial steady state difference.

1.2.2.5 Distinct pathways are involved in transcriptional/post-transcriptional control of RAG constitutive and inducible mRNA expression

What controls the alterations in RAG mRNA expression mediated via altered transcription rates or mRNA stability? Two obvious possibilities are de novo synthesis of repressor protein(s), or alternatively, post-translational modifications such as phosphorylation of trans factors. The molecular mechanisms involved in the signal transduction-mediated transcriptional/post-transcriptional alterations of RAG transcripts have only just begun to be assessed, and the picture is complex. I have previously found that the antigen-receptor signaling-mediated increase in RAG expression is dependent upon de novo protein synthesis. Pre-treatment with the protein synthesis inhibitor cyclohexamide prior to PMA+ionomycin stimulation abrogated increases in both RAG1/2 transcription rates and total RNA
levels [75]. Furthermore, Takahama and Singer reported that RAG mRNA downregulation observed upon TCR-β cross-linking or PMA treatment of CD4<sup>+</sup>CD8<sup>+</sup> fetal thymocytes was reversed by cycloheximide co-incubation [221]. A requirement for newly-synthesized proteins has also been observed in RAG protein downregulation in DN and DP thymocytes induced by stimulation with PMA+ionomycin [218].

While the above reports argue for a de novo protein synthesis-dependent pathway of RAG mRNA regulation, RAG downregulation is not abrogated by pretreatment with cyclohexamide in the pre-B cell line REH and the pre-T cell line CEM [218]. This implies that this mechanism occurs via a post-translational pathway, such as tyrosine phosphorylation-mediated activation (or repression) of a constitutively-expressed PTK. In this context, v-abl has a de novo-independent negative regulatory effect on RAG1/2 mRNA and protein expression in a temperature sensitive A-MuLV transformed pre-B cell line [220]. Specifically, the activation of v-abl at the permissive temperature in this line shuts off RAG1/2 mRNA whereas v-abl inactivation at the non-permissive temperature results in RAG mRNA upregulation. It remains to be determined, however, if this finding is physiologically relevant since v-abl is is not a protein expressed in normal cells.

The different mechanisms described in different reports correlate well with the differing kinetics of the observed effects. For example, the PMA/ ionomycin-induced upregulation of RAG mRNA in OCI-LY8 does not peak until 6 h [75], while protein synthesis-independent alterations in RAG expression in other reports, via either cAMP or PKC-dependent pathways were observed to peak within 1h [218, 220]. This is consistent with the rapid kinetics of phosphorylation-regulated phenomena such as the phosphorylation-dependent activation of NF-κB in pre-B cells by PMA, which is measured in minutes rather than hours [283].

Other features of RAG regulation apart from BCR or TCR signaling-mediated alterations may also require de novo protein synthesis. For example, cycloheximide treatment of thymocytes alone can increase constitutive RAG expression 6-8 fold without affecting expression of other genes such as c-fos, c-myc, or TCR-β. This suggests the presence of a rapidly-degraded RAG-specific regulatory element constitutively expressed in pre-T cells that has to be continually synthesized for maintenance of
steady state expression [218]. Furthermore, in contrast to the de novo independent, signaling pathway of RAG mRNA destabilization/decreased transcription shown by Neale et al., treatment of cell lines with cycloheximide alone causes 4-6 fold increases in RAG1/2 transcripts. This suggests at least two repressor pathways in this system: one in which RAG transcription is under the control of a de novo synthesized repressor and a second which is PMA-inducible and likely post-translationally modified [218]. Finally, at least one of the regulatory proteins responsible for mediating constitutive increases in RAG1 expression in the RAG\textsuperscript{bi} OCI LY8 variants may be newly-synthesized in culture, as suggested by the observation that protein synthesis upregulates RAG expression in the RAG\textsuperscript{bo} parental clone OCI LY8-C3P [75]. The fact that this increase takes place largely in OCI LY8-C3P suggests that the parental clone constitutively expresses a negative regulatory element at higher levels than in C3-A11N. This selective increase cannot be attributed to the possibility that RAG mRNA expression is already maximal in C3-A11N because RAG expression can be upregulated by PMA+ionomycin stimulation in this variant. Because cycloheximide treatment does not appear to upregulate OCI LY8-C3P RAG expression to levels observed in the RAG\textsuperscript{bi} variant C3-A11N, this suggests that this de novo synthesized factor does not completely account for the differential expression of RAG in these variants. Therefore, other negative regulatory factors in OCI LY8-C3P (or positive regulatory factors in C3-A11N) that are not synthesized de novo may also be present. Nevertheless, the fact that constitutive expression increases via cycloheximide treatment in several different systems suggests that there is a transcriptional repressor present throughout lymphocyte development with a relatively short \( t_{1/2} \). However, cycloheximide treatment does not result in RAG expression in RAG mature B cells or non-lymphoid cells, suggesting a requirement for active transcription of the RAGs [75], an idea consistent with RAG tissue-specific repressor/silencer elements.

1.2.3 RAG regulation is only one level regulating V(D)J recombination

Tissue and developmental-specific regulation of the only lymphoid-specific and essential recombinase components, the RAGs, is perhaps the most obvious level of controlling the V(D)J recombination reaction, and as a major focus of this thesis, has already been extensively discussed. However, it is important to point out that RAG regulation is not the only regulatory mechanism controlling V(D)J
recombination. Evidence of additional locus-specific regulatory mechanisms comes from observations in several different contexts. First, non-lymphoid cells transfected with the RAGs, and that rearrange artificial recombination substrates, do not rearrange their endogenous antigen receptor loci (referred to as “inter-lymphoid specificity”; [284, 285]). Secondly, while all antigen receptor rearrangements use the same RSS elements and recombinase machinery, the rearrangement of Ig or TCR genes is largely restricted to B and T cells, respectively (“intra-lymphoid” specificity; [286]). Thirdly, rearrangements at Ig and TCR loci are not initiated simultaneously upon RAG1/2 expression, but are generally rearranged in a set temporal pattern, both within and between different antigen receptor loci (“developmental specificity”; reviewed in [287]). For example, with respect to rearrangement of Ig loci throughout B cell development, generally, D_{H}-J_{H} is followed by V_{H}-DJ_{H} joining followed by V-J joining at the LC loci). Finally, it is well documented that some V_{H}, D_{H}, and J_{H} gene segments are rearranged at higher frequencies than others (reviewed in [288, 289]). Although it can be argued that this results from antigenic selection, this cannot always serve as an explanation. For example, certain pseudogene segments not coding for a surface IgH protein are still preferentially rearranged [290, 291].

Three major RAG-independent factors that regulate the efficiency of V(D)J recombination at various antigen receptor loci are: RSS quality, coding end microhomologies, and chromosomal accessibility. Other factors that can also determine usage of various antigen receptor loci over others are cellular selection and genetic differences (reviewed in [292]). Also, some of the potential binding factors not deemed essential for recombination (discussed earlier in section) may indirectly regulate RAG activity by competing with the RAGs in a lineage- and/or stage-specific manner for RSS binding, and as such may also be regulators of the reaction. Finally, lineage- and stage-specific transcription factors regulating antigen receptor loci enhancers also play a role in regulating recombination. These and others involved in the regulation of B cell specific genes are beyond the scope of this thesis and will not be detailed further in this section.
1.3 Differential expression analysis: an approach for identifying novel B lineage developmental genes

1.3.1 General considerations for isolating differentially expressed genes

The human genome encodes ~100,000 genes, about half of these already available as expressed sequence tags (ESTs) [293]. A single cell expresses only 10-20,000 of these genes, with approximately 95% being rare message classes (<50 copies/cell). In terms of overall abundance, however, these are largely outnumbered by the small fraction of ubiquitously-expressed, housekeeping gene transcripts. Developmentally restricted genes, i.e. differentially expressed within the developmental pathway of a given cell lineage (including B cell differentiation), largely belong to the rare message class. Over the years, several approaches have been developed for the isolation of differentially expressed genes. These techniques can be subdivided into two categories: non-PCR based and PCR-based, which can each be further subdivided into subtractive and comparative. In the following sections, the essential features of these approaches will be discussed. A summary of their relative strengths and weaknesses is also outlined in table 1.6.

1.3.2 Classic (non PCR-based) methods for isolating differentially expressed genes

1.3.2.1 Differential screening

In differential screening, radiolabeled cDNA probes are generated from poly(A)+ RNA from two cell types of interest and hybridized separately to duplicate filters of a cDNA library [294]. Clones hybridizing to both probes correspond to genes commonly expressed by both cell types, whereas clones hybridizing to only one probe represent genes uniquely present in one cell type. This is an easy and reliable technique, and has led to the isolation of numerous genes. A few examples include genes temporally regulated during development of Drosophila [295], genes induced by drug treatment [296], and virus transformation [297], and genes related to cell cycle [298]. However, this method works well only for abundantly-expressed genes since the cDNA probes obtained from the entire population of mRNAs are of very high complexity and the representation of rare mRNAs is extremely low. The approximate limitations of detection are mRNAs representing 0.1% of the total population [299].
| Technique | Basic principle | Detection limit (x) | Sensitivity | Specificity | Selectivity | Linearity | Reproducibility | Robustness | Compatibility | Amplification | Relative sensitivity | Relative selectivity | Relative specificity | Relative linearity | Relative reproducibility | Relative robustness | Relative compatibility |
|-----------|----------------|-------------------|-----------|-----------|-----------|---------|------------|---------|-----------|-------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| Negative Selection | Based on specific binding of complementary nucleic acids to a matrix followed by elution and washing, and then extracting the nucleic acids | 10-8 to 10-11 | High sensitivity | High specificity | High selectivity | Good linearity | Good reproducibility | Good robustness | Good compatibility | PCR amplification | High sensitivity | High selectivity | Good linearity | Good reproducibility | Good robustness | Good compatibility |
| Positive Selection | Based on specific binding of complementary nucleic acids to a matrix followed by elution and washing, and then extracting the nucleic acids | 10-12 to 10-14 | Low sensitivity | Low specificity | Low selectivity | Poor linearity | Poor reproducibility | Poor robustness | Poor compatibility | PCR amplification | Low sensitivity | Low selectivity | Poor linearity | Poor reproducibility | Poor robustness | Poor compatibility |
| Competitive Hybridization | Based on the hybridization of specific nucleic acids to a matrix followed by elution and washing, and then extracting the nucleic acids | 10-10 to 10-12 | Moderate sensitivity | Moderate specificity | Moderate selectivity | Moderately good linearity | Moderately good reproducibility | Moderately good robustness | Moderately good compatibility | PCR amplification | Moderate sensitivity | Moderate selectivity | Moderate specificity | Moderately good linearity | Moderately good reproducibility | Moderately good robustness | Moderately good compatibility |
| Northern Blotting | Based on the hybridization of specific nucleic acids to a matrix followed by elution and washing, and then extracting the nucleic acids | 10-14 to 10-16 | Very low sensitivity | Very low specificity | Very low selectivity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility | PCR amplification | Very low sensitivity | Very low selectivity | Very low specificity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility |
| Southern Blotting | Based on the hybridization of specific nucleic acids to a matrix followed by elution and washing, and then extracting the nucleic acids | 10-16 to 10-18 | Very low sensitivity | Very low specificity | Very low selectivity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility | PCR amplification | Very low sensitivity | Very low selectivity | Very low specificity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility |
| In Situ Hybridization | Based on the hybridization of specific nucleic acids to a matrix followed by elution and washing, and then extracting the nucleic acids | 10-18 to 10-20 | Very low sensitivity | Very low specificity | Very low selectivity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility | PCR amplification | Very low sensitivity | Very low selectivity | Very low specificity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility |
| Microarray | Based on the hybridization of specific nucleic acids to a matrix followed by elution and washing, and then extracting the nucleic acids | 10-20 to 10-22 | Very low sensitivity | Very low specificity | Very low selectivity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility | PCR amplification | Very low sensitivity | Very low selectivity | Very low specificity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility |
| Quantitative PCR | Based on the hybridization of specific nucleic acids to a matrix followed by elution and washing, and then extracting the nucleic acids | 10-22 to 10-24 | Very low sensitivity | Very low specificity | Very low selectivity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility | PCR amplification | Very low sensitivity | Very low selectivity | Very low specificity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility |
| Real-Time PCR | Based on the hybridization of specific nucleic acids to a matrix followed by elution and washing, and then extracting the nucleic acids | 10-24 to 10-26 | Very low sensitivity | Very low specificity | Very low selectivity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility | PCR amplification | Very low sensitivity | Very low selectivity | Very low specificity | Very poor linearity | Very poor reproducibility | Very poor robustness | Very poor compatibility |

Notes:
1. Advantages and limitations are relative to other comparative techniques with each other.
2. Sensitivity, specificity, selectivity, linearity, reproducibility, and robustness are relative to each other for more general differences between comparative and subtractive, PCR-based and non-PCR-based, and quantitative and qualitative techniques.
1.3.2.2 Microarray analysis

Microarray analysis of cDNA clones is a recent differential screening technique done at a microscopic level [300-304]. This technique takes advantage of the current availability of human and mouse cDNA sets in databanks. In other words, at least partial cDNA sequences of genes (ESTs) must be available. Microarrays are generated by using cDNA clones or 3' end portions of clones (usually non-redundant 20 bp oligonucleotides) whose full or partial sequences are known (either ESTs or full-length cDNAs). Set amounts of these clones are added in duplicate (usually with one of oligonucleotides having a 1 bp mismatch to control for reproducibility of the arraying and hybridization) in microtitre plates. Samples from these plates are then robotically printed (dot blotted) onto glass "chips" at very high density in very small areas (up to 20,000 cDNAs can be simultaneously monitored/chip). Fluorescent probes are then prepared from mRNA from the sources being compared by RT. As with differential screening, these complex cDNA mixes are then used as probes, except in the case of microarray analysis, the "chips" rather than cDNA libraries, are probed. This is done in very small volumes with appropriate quantitation standards to control for specificity and other controls for non-specific binding, including microarraying known housekeeping genes, to assess variation. Originally, microarray hybridizations were done independently with duplicate microarrays, but recently an attractive modification has been described in which both fluoresceinated probes are mixed and differential expression is by means of simultaneous, two-color fluorescence hybridization [301]. In this report, cDNA from one population (in this case a transgenic cell line) is fluoresceinated with a red dye, and the second population (wild type) is fluoresceinated with green dye. If approximately equal levels of two genes are present, microarray dots appear as yellow or brown. If differential expression is seen in either the transgenic or wild type cells, the dots appear as either red or green respectively. The technique is currently limited by the sizes of databank sets of sequence information and by the sophistication of algorithms required for large-scale analysis. However, as the nucleotide sequence of the entire genome of several organisms is uncovered, it may be possible to screen all known genes for differential expression in parallel. Therefore, once all genes have at least partial corresponding cDNAs, this may be the most powerful and efficient way of assessing differential gene expression. It must also be kept in mind however, that different isoforms of genes may be present, based on lineage.
and developmental stage. In this context, microarray may not only depend on the availability of complete set of cDNAs, but on new techniques, such as laser capture microdissection, which allows isolation of specific cells from heterogeneous populations and analysing minute amounts of RNA present in isolated cells. Microarray analysis may open various avenues; for example powerful correlations can be directly made between human gene sequence information and various disease situations and diagnostics in clinical medicine will be greatly improved.

1.3.2.3 Subtractive hybridization

To clone differentially expressed rare mRNAs, subtractive hybridization provides a powerful alternative (reviewed in [299, 305]). As shown in Figure 1.12, the technique uses a process called driver excess hybridization. Briefly, mRNA from which one wants to isolate differentially expressed sequences is reverse transcribed into cDNA (termed the tester). The cDNA is then hybridized iteratively to an excess (≥10 fold) of mRNA from cell type B (termed the driver). cDNAs corresponding to mRNAs expressed in both cell types will form DNA: RNA hybrids while cDNAs corresponding to mRNAs uniquely expressed by cell A remain single-stranded. The unhybridized cDNAs are then separated from the DNA: RNA hybrids and excess driver by chromatography on hydroxyapatite columns or via biotinylation (for review on variations on subtractive hybridization and related techniques see [305]). This is the subtraction step. These cDNAs can then be used either as probes to screen a cDNA library or as templates to synthesize double-stranded cDNAs for the construction of a subtracted cDNA library. This approach has led to the discovery of such important lymphocytic genes as those encoding the TCR [306], and λ5 [307], and is still widely used in the search of genes differentially expressed in different cell types or different cellular states. Unlike differential screening, this technique is effective for the detection and concentration of rarer mRNA species. This technique, however, is time consuming and exhibits low reproducibility. Additionally, it is not efficient for isolating genes for which the expression level is only a few times higher in the positive relative to the negative cells (quantitative differences) because mRNA from negative cells is always used in a large excess in subtractive hybridization. Finally, a technical limitation to its application is the large amount of starting material required to drive hybridization to completion. One
Figure 1.12. General outline of subtractive hybridization. Two samples are mixed together (driver sequences are present in excess), denatured, and allowed to anneal. Duplexes formed between driver and tracer (the latter also referred to as the tester; denoted by asterisks) are then removed, as is unhybridized driver, leaving a population enriched for sequences present in the tracer but absent in the driver. Different sequences are indicated by varied colors. Red sequences are differentially expressed target sequences, present in the tracer (tester) populations but absent (or present at lower levels) in the driver cDNA population. For simplicity, only one differentially expressed tester species is shown in red, but in fact there are several, if not many possible differentially expressed sequences. Additionally, relative transcript abundance variation is not shown. Depending on the type of subtraction, the target and driver populations can be comprised of mRNA, single or double-stranded cDNAs, or even entire cDNA libraries (for details see review by Sagerstrom et al.; reference [305]).
solution to this problem is amplification of mRNA from a single or a few cells by PCR [308].

1.3.3 PCR-based approaches for identifying differentially expressed genes

1.3.3.1 Differential Display (DD RT-PCR) and the basic principle behind standard DD RT-PCR methodology

Differential display is a relatively new methodology (first developed by Liang and Pardee in 1992) for detecting genes that are differentially expressed under certain conditions or between different cell subsets or lines [309, 310]. The basic strategy of this method involves several steps, as schematically represented in Figure 1.13. First, poly (A)+ RNAs are extracted from the various cell sources being compared. The second step involves subdividing (fractionating) mRNAs from each source into subsets or "pools" by reverse transcription using modified (anchored) (dT) oligonucleotides. Next, partial 3' end cDNA sequences from subsets of RNAs are amplified under low stringency PCR conditions (40-42°C annealing cycles) in the presence of 35S-radiolabeled dNTPs using the same primer used for RT in combination with various arbitrary decamers.

The design of the primers used in the RT and PCR (steps 2 and 3, respectively) is the key theoretical aspect of this technique. The 3' primers make use of the poly (A) tail present on most eukaryotic RNAs plus two additional 3' bases. These two 3' bases add partial specificity to the amplification reaction since they reverse transcribe only one twelfth of the RNAs i.e. 12 different combinations of the last two 3' bases (omitting T as the penultimate base). Random primers are designed for the 5' end and it has been estimated that 20 such random 10-mers upstream of three possible anchored oligo (dT) primers statistically cover the 15,000 estimated mRNA sequences within a given cell while at the same time generating an ideal amount (150-200) of these species per primer pair for "display". A series of experiments showed that the pattern of amplification products with any set of primers yields highly reproducible patterns in independent amplifications.

The fourth step is the "display" portion of the procedure, and involves running the 150-200 short (100-600 bp) radiolabeled amplicon products from different cell types (generated by each primer pair in the RT-PCR step) on denaturing sequencing gels side by side for comparison. In the final step, cDNAs of interest are selected, recovered by excision from the dried sequencing gel, and subsequently
The principle of differential display RT-PCR (DD RT-PCR). Total cellular RNA is extracted from the sources being compared, in the case of the work in this thesis, these are OCI LY8-C3P (RAG<sup>89</sup>) and C3-A11N (RAG<sup>56</sup>). The samples are then treated with DNase I to remove cellular DNA contamination and the RNA is subdivided into cDNA subsets by reverse transcription with several distinct mRNA-targeting oligo (dT) primers having two anchoring nucleotides to initiate reverse transcription at the start of the eukaryotic poly (A) tail (see text). The identical oligo (dT) and an arbitrary 10-mer upstream oligo are used in a subsequent low-stringency PCR to amplify a specific pattern of DNA bands. The display of patterns between OCI LY8-C3P and C3-A11N are compared and differences represent the 3' ends of potentially regulated genes. Bands of interest are then reamplified and purified.
eluted via precipitation. The criteria for selection is differential expression of banding, hence bands appearing in one cell type but not the other are considered to be derived from differentially expressed genes. The eluted products are then reamplified under similar conditions as the original PCR step and purified from high-percentage agarose gels for subsequent manipulations, such as subcloning and sequencing, or for generating probes for further screening.

The validity of the technique was determined by assessing whether an amplicon detected as differentially expressed was truly differentially expressed between cell lines or differing conditions. In these experiments, reproducible bands unique to one variant were excised, purified, reamplified and then used as probes in Northern analysis of the same total RNA as the batch used for differential display. The results of these Northern blotting experiments demonstrated that only a very small fraction of differentially expressed bands could detect differentially expressed transcripts with the initial protocol [309]. Therefore, Northern blot reconfirmation became particularly important as a screening step since the majority of DD amplicons either represented false negatives or positives i.e. the same or no signal detected by Northern blot, respectively. In this context, the term "differential display" originally referred to just the RT, PCR, and display steps, but is now commonly used to refer to the entire screening approach including the subsequent Northern blot screens. To increase the efficiency of the methodology, Liang and Pardee tested various parameters of the original technique and subsequently streamlined it in 1993 [311]. The key modifications were: 1) using total RNA instead of polyA RNA to decrease smearing (since polyA RNA extraction uses oligo-dT separation which can act as mispriming contaminants for the anchored primer), 2) treating samples with DNAse to prevent genomic DNA contamination in subsequent steps, 3) running multiple, independent samples in both RT and PCR steps to increase reproducibility of banding patterns, 4) running -RT controls in parallel for each source to eliminate spurious banding, 5) reducing redundancy in the technique by reducing the number of anchored primers from 12 to just four degenerate primers. Although these basic modifications are now accepted parts of standard DD RT-PCR methodology, it still appears that in most situations, only 10-15% of differentially expressed bands can be reconfirmed by Northern blotting. Efforts to further streamline the technique will be detailed in the general discussion.
1.3.3.2 Other comparative DD RT-PCR-like techniques

Recently, similar alternative methods to DD RT-PCR for comparing gene expression profiles (aka. RNA fingerprinting) have been described (see Table 1.7). The first, RAP-PCR, for RNA Arbitrary Primed PCR, is conceptually very similar to DD RT-PCR; the key difference is that long (20 bp) completely arbitrary primers are used in both the reverse transcription and PCR step in place of the series of anchored, degenerate dT primers [312]. This results in much larger products being generated which are not necessarily biased for the polyadenylated regions of genes. Therefore, one inherent advantage of this technique over DD RT-PCR is that a more informative sequence is obtained directly from clones: often the entire coding regions can be potentially isolated directly without library screening or 5' RACE [313]. Furthermore, non-eukaryotic RNAs can be analyzed by this technique (such as certain bacterial RNAs). Additionally, the length of the primers used in the technique potentially increase reproducibility relative to DD RT-PCR as higher stringency amplification conditions can be tolerated, thus decreasing the amount of mispriming. However, on the downside, RAP-PCR, unlike DD RT-PCR, can also select for other RNAs besides mRNAs [314]. This is potentially a major problem since ribosomal and transfer RNA species comprise the predominant (95%) eukaryotic RNA species [315]. Additionally, because only single primers are used, only palindromic RNA sequences may be detected. Recently, RAP-PCR has been modified by using two or more arbitrary primers to minimize this problem [316]. Because of the large size range of RAP-PCR products (generally up to 4 kb), cDNAs are resolved on agarose rather than sequencing gels. Although the use of agarose gels is more straightforward relative to sequencing gels, it severely limits the resolution to only a few distinct amplicon products. Another limitation with RAP-PCR is that the large length of the primer and its completely arbitrary nature means that there is no systematic way of detecting most of the mRNAs, unlike DD RT-PCR where mRNAs are subdivided into subsets and are further subdivided systematically via PCR. This results in two major technical difficulties: redundancy and a large number of primer sets having to be employed.

The second technique, RLCS (Restriction Landmark cDNA Scanning), is conceptually quite different from both DD RT-PCR and RAP-PCR [317]. The basic principle of this technique is to use restriction enzymes as landmarks so that cDNAs can be displayed as “spots” via high-resolution two-dimensional
## Table 1.6

Summary of comparative techniques for identification of differentially expressed (RNA fingerprinting approaches)

<table>
<thead>
<tr>
<th>Standard Technique/ original reference</th>
<th>Comparison of basic steps comprising procedures</th>
<th>Relative advantages(^1)</th>
<th>Relative limitations</th>
</tr>
</thead>
</table>
| Differential display (DD RT-PCR)       | 1. RT: subdivision of total RNAs with series of anchored (dT)\(_n\) primers  
2. PCR amplification with same primer plus one arbitrary primer and 32P-dNTPs  
3. Separation of cDNA banding patterns on denaturing sequencing gel and autoradiography  
4. Excision of 100 bp-600 bp differentially expressed cDNAs; PCR reamplification under same conditions and primers  
5. Separation of reamplified products on agarose gel and elution for cloning, sequencing, etc. | * simultaneous comparison of multiple cell lines or conditions  
* specific for mRNAs  
* relatively few primer combinations required in PCR (partially specific anchoring primer in RT and PCR subdivides cDNAs into subsets) | * shortening of 3' sequences-limited information from A-rich regions  
* false positives due to low stringency PCR conditions  
* low reproducibility- short primers |
| RNA Arbitrarily Primed PCR (RAP-PCR)   | 1. RT: total RNA with one arbitrary primer  
2. PCR amplification with same primer and 32P-dNTPs  
3. Separation on DNA sequencing gel and autoradiography  
4. Excision of 100-600 bp differentially expressed cDNAs; PCR reamplification with 32P-dNTPs  
5. Separation of reamplified products on agarose gel and elution for cloning, sequencing, etc. | * simultaneous comparison of multiple cell lines or conditions  
* more immediate informative sequence information: can amplify large cDNAs, open reading frames; not biased for poly(A) regions  
* only one primer required for RT for all reactions  
* longer primers and higher stringency PCR conditions- increased reproducibility | * subject to increased amount of abundant contaminant products: amplification of non-specific ribosomal and transfer RNAs;  
* very large amount of primer combinations required in PCR (due to degeneracy of primer in RT and PCR step)  
* no systematic method for covering all mRNAs |
| Restriction Landmark cDNA Scanning (RLCS) | 1. RT: poly(A)+ RNA with biotinylated, anchored (dT)n primer  
2. Generation of 1st "landmarks" i.e. protruding cohesive 5' terminal via complete digestion of restriction sites 5' to primer with enzyme A  
3. Labelling of digestion products with 32P-dNTPs; purification via streptavidin beads  
4. Electrophoresis in 1st dimension  
5. Generation of 2nd "landmarks" i.e. independent protruding cohesive 5' terminal via complete digestion by of restriction sites 5' to primer with enzyme B  
6. Electrophoresis in 2nd dimenion  
7. Excision of 100-2000 bp cDNA "spots" from two dimensional gels; ligation into vectors, cloning via PCR amplification using vector-specific primers  
8. Separation of products on agarose gel and elution for sequencing | * larger 3' sequences can be generated and resolved: up to 2 kb  
* greater separation potential in two dimensions- may lower # co-migrating contaminating cDNAs  
* correlation between relative abundance seen on RLCS gels and reconfirming northerns  
* analysis of 10X more products/gel  
* no false positives resulting from low stringency PCR conditions: no PCR amplification required for visualization of spots | * decreased practicality of comparing >2 cell lines or conditions  
* false positives or negatives from incomplete digestion with either enzyme A or B |

\(^1\) ? denotes potential/theoretical advantages but that have not yet been empirically established

84
Gened

Introduction
gel electrophoresis (see Table 1.7 and Figure 1.14). The major advantage of this technique over other comparative techniques is that large portions of mRNAs can be systematically analyzed on a relatively small number of gels. Additionally, as a non-PCR based technique, false positives that arise from amplification conditions are not an issue. As with RAP-PCR, this technique has an advantage over DD RT-PCR in that larger products can be detected, hence, allowing informative sequence information to be immediately obtained. However, false positives can result from other sources in this technique such as incomplete digestion with either restriction enzyme. Additionally, unlike DD RT-PCR and RAP-PCR, it is impractical to compare more than two cell lines or conditions at a time.

In addition, various useful variations on the DD RT-PCR technique (or the closely related RAP-PCR technique) have also been suggested for situations in which one already has structural information on the genes one is interested in identifying. One of the most interesting of these is to bias sampling towards particular motifs that are conserved across genes using "motif-specific" primers in place of arbitrary or anchored primers. Using this modified methodology, various successes in cloning novel multigene family members have been reported, such as the use of SH2 or SH3 domain primers to identify novel protein kinases, or the use of zinc finger primers for identification of novel transcription factors [313].

1.3.3.3 Representational Difference Analysis of cDNA (cDNA-RDA)
cDNA-RDA was introduced as an alternative PCR-based technique to DD RT-PCR and has the potential to rapidly reduce the number of candidate genes to a few which could be easily characterized [318]. The cDNA-RDA technique couples a subtractive hybridization step with PCR amplification and was originally developed for use with genomic DNA as a method for isolating the difference between two complex genomes [319]. The original RDA technique has led to several key discoveries including probes for the detection of genetic lesions, sequences from several unknown pathogens, and isolation of polymorphic markers linked to disease traits (reviewed in [320]). In contrast to DD RT-PCR, cDNA-RDA eliminates by a subtractive hybridization step those fragments present in both populations, selecting for the differences. Similar techniques based on the combination of subtractive hybridization and PCR have also been described, such as suppression subtractive hybridization (SSH) [321].
Figure 1.14. Principle of Restriction Landmark cDNA Scanning (RLCS) schematically represented. A. Preparation of the RLCS sample. The red line (a) and the double blue lines (b-f) indicate poly(A)+ RNA and double-stranded cDNA, respectively. Asterisks show the radiolabeled nucleotides incorporated. An oligo(dT) anchor primer with MA at the 3' -end and \([\alpha-^{32}P]dGTP\) are used for cDNA synthesis and labeling, respectively. B. Separation of cDNA species by two-dimensional gel electrophoresis. cDNA fragments in an RLCS sample (CS1, CS2, and CS3) are separated by two-dimensional gel electrophoresis and detected as spots CS1, CS2, and CS3 respectively. The figure is adapted from Suzuki et al., 1996 (reference [317]).
linker capture subtraction (LCS) [322], and selective amplification via biotin-and restriction-mediated enrichment (SABRE) [323]. The two most commonly used PCR based subtractive methodologies, cDNA-RDA and SSH, are outlined in more detail in Figure 1.15.

1.3.3.4 Serial Analysis of Gene Expression (SAGE)

SAGE is another PCR-based method that has the potential to identify differentially expressed genes [324, 325]. SAGE is based on two principles. First, short nucleotide sequence tags (9-10 bp) from the complete mRNA population each contain sufficient information to uniquely identify a transcript, provided it is isolated from a defined position within the transcript. In other words, a sequence of 9 bp can distinguish 262, 144 transcripts \((4^9)\) given a random nucleotide distribution, far more than the 80,000 estimated human transcripts. Second, concatenation of short sequence tags allows the efficient analysis of transcripts in a serial manner by the sequencing of multiple tags within a single clone. By sequencing a large number of these concatenated, tagged clones, the presence and abundance of individual transcripts can be determined. Comparison of the information obtained from two cell types will reveal both qualitative and quantitative differences. Upon identification of such differences, the tag sequence provides the required information to reliably match with known genes and databank ESTs. As with microarray analysis, this technique relies on the availability of databank cDNA information. The general scheme of this approach is detailed in Figure 1.16.

1.3.4 Various ways DD RT-PCR or other differential expression techniques have been used to identify novel lymphocyte-specific genes and/or genes involved in lymphopoiesis

Of the ~400 novel human and mouse genes cloned by differential expression approaches in last three years, >200 have been identified by DD RT-PCR alone. These include novel genes that are lymphocyte specific and/or have a role in B and/or T lymphopoiesis. Examples of these, the majority of which have been cloned by DD RT-PCR or cDNA-RDA, are shown in Table 1.5.

How have DD RT-PCR and other differential expression analysis methods contributed to our understanding of molecular mechanisms underlying lymphopoiesis? First, they have been very useful in the identification of novel downstream targets of signaling pathways and transcription factors. With respect to downstream targets of signaling, DD RT-PCR has been particularly effective for identifying
Figure 1.15. General schemes of the PCR-based subtraction methodologies known as cDNA Representational Difference Analysis (cDNA-RDA) and Suppression Subtractive Hybridization (SSH). Like with classical subtractive hybridization, all PCR-based subtractive schemes use two populations: a tester (aka tracer) population and a driver population, used to subtract common target sequences. For both cDNA-RDA and SSH, tester and driver double stranded (ds) cDNAs are first digested with a four base-cutting restriction enzyme that yields blunt ends. For cDNA-RDA, the procedure involves repeated rounds of subtraction, each round which can be divided further into two phases: 1) generation of amplicons and 2) generation of difference products. In the primary phase (generation of amplicons), both tester and driver samples are ligated with the same unphosphorylated adapters, mixed, denatured and hybridized, and subjected to end-filling and PCR amplification. After amplification, both driver and tester amplicons are digested with the same restriction enzyme to cleave away adapters. During the second phase (generation of representation products), the tester amplicons were ligated to new adapters and mixed with excess unligated driver, heat-denatured, and hybridized. The mixture is then subjected to end filling, mung bean nuclease digestion (to remove ssDNAs including excess unhybridized driver) and PCR amplification of hetero and homo-hybrids. This time, only the differentially expressed tester cDNA homo-hybrids can be exponentially amplified, resulting in a dramatic enrichment in differentially expressed sequences (referred to as primary representational difference products, or RD1). RD1 can then be re-introduced at the beginning of the procedure and the whole process after being repeated 2-3 times, results in further enriched difference products i.e. RD2, RD3, etc.... The process is shown up to the first difference product. Although only four populations are shown for simplicity, one round shows complete enrichment, in reality, the tester with some common sequences still left are reintroduced at the stage indicated for further enrichment i.e. to generate second and third difference products, products are reintroduced into the scheme at the tester stage in proportions of driver excess. For SSH, the tester cDNA fragments are divided into two portions ligated with two different adapters, resulting into two tester populations. Additionally, unlike cDNA-RDA, SSH uses only a single round of subtraction, but two distinct hybridization steps within this round. In the first hybridization step, an excess of driver without adapter is individually added to each adapter-ligated tester sample. After samples are allowed to heat-denature and reanneal, differentially-expressed cDNAs are enriched via elimination of common non-target cDNA by hybrid formation with driver sequences, and like with second-order kinetics seen in classical subtractive hybridization [305], a normalization (equalization) of high and low-abundance ss cDNAs occurs. During the second hybridization, the two primary hybridization sequences are mixed together without heat-denaturing, thereby allowing only ssDNA tester sequences to re-anneal. A portion of excess, denatured driver is added to further enrich for tester hybrids. The total population of hybridized DNA is then subjected to two rounds of PCR to amplify the desired target sequences. In the primary PCR, the tester inter-population heterohybrids have the important feature of having the two distinct adapter sequences at their 5' ends, distinguishing them from intra-tester homohybrids, which have only adapters from one or the other original tester populations. Therefore, in addition to species with no adapters (incapable of being amplified) or with one adapter (capable of being only linearly amplified; these two cases also seen in cDNA-RDA), the homohybrid tester species (with identical adapters on both their ends) form self-annealing panhandle structures, and cannot serve as templates for PCR; hence the "suppression" PCR effect. The result is that the two adapters on tester heterohybrids allow selective exponential PCR amplification of tester hybrids from only the two independently normalized, subtracted populations. Finally, in the second PCR, nested primers are used to reduce any background while further enriching the differentially expressed sequences. For cDNA-RDA, solid boxes represent the oligonucleotides used to generate the representations and hatched boxes show the oligonucleotides used to generate difference products. With SSH, these two box types represent the two different linker types used to subdivide tester populations. Figure based on the methodologies outlined by Liseisyn et al., 1993, Hubank et al., 1994, and Diatchenko et al., 1996 (references [318, 319, 321]).
Figure 1.16. The principle of Serial Analysis of Gene Expression (SAGE) schematically represented. Double stranded cDNA is synthesized from mRNA using biotinylated oligo (dT) primers. The cDNA is then cleaved with restriction endonuclease having a 4 bp-cutting recognition site (anchoring enzyme) that cleaves most transcripts at least once (1 in $4^4$ or every 256 bp). The most 3' portion of the cleaved cDNA is isolated by streptavidin bead purification, thereby creating a unique site on each transcript which corresponds to the restriction site closest to the poly (A) tail. The cDNAs are then divided in half and ligated via the anchoring restriction site to one of two linkers containing a type IIS restriction site (tagging enzyme). Type IIS restriction endonucleases cleave at a defined distance up to 20 bp away from their asymmetric recognition sites. The linkers are designed such that cleavage of the ligation products with the tagging enzyme results in the linker with a short piece of cDNA sequence (tag). The two pools of the released tags are ligated to each other. Subsequent PCR amplification of the ligated tags with primers specific to each linker selects ligation products linked tail to tail (ditag). The linker sequence is removed from the PCR products with the anchoring enzyme. The ditags are concatenated and cloned. A library is thereby created in which each mRNA transcript is represented by a short tag. By sequencing a large number of these concatenated ditag clones, the presence and abundance of individual transcripts can be determined. Comparison of the information obtained from two cell types will reveal both qualitative and quantitative differences. Upon identification of such differences, the tag sequence plus the anchoring enzyme can provide enough information for isolation of the corresponding full-length cDNA clones because this is enough information to reliably match with databank EST or full length cDNA sequences. Sequences that are colored red and green represent primer-derived sequences, whereas blue represents transcript-derived sequences, with X and O indicating nucleotides of different tags. The figure is adapted from Velculescu et al., 1995 (reference 1324).
<table>
<thead>
<tr>
<th>Gene</th>
<th>Approach used for original isolation</th>
<th>General strategy/cell lines/tissues used for subtractions and/or comparisons in original screens</th>
<th>Reported expression pattern</th>
<th>Postulated/characterized function in lymphocyte development</th>
<th>General structural classification</th>
<th>Year identified/ original reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCR β chain</td>
<td>Subtractive hybridization</td>
<td>Subtracted probes made from Th mouse hybridoma tester and B cell hybridoma driver; used to screen a subtracted library (Thβ) constructed from another Thβ hybridoma/b cell combination</td>
<td>T cell-specific; individual components restricted to various thymocyte and T cell subsets</td>
<td>T cell antigen-specific recognition of MHC class II restricted T cells; helper activity (stabilizes TCR-MHC class I interaction)</td>
<td>Ig superfamily (SF) glycoprotein which are components of the TCR antigen receptor complex; heterodimer of α (45 kDa) and β (40 kDa) or γ (45 kDa) and δ (40 kDa) chains disulfide linked. Noncovalently associated to the CD3 chain</td>
<td>Hedrick et al., 1984; [306]</td>
</tr>
<tr>
<td>TCR γ chain</td>
<td>Subtractive hybridization and Differential screening</td>
<td>Subtracted library made from two closely related Th hybridomas; differentially screened with two probes: subtracted probe generated from Th γ hybridoma vs A20 B lymphoma, and A20-specific probe</td>
<td>T cell-restricted; individual components restricted to various thymocyte and T cell subsets</td>
<td>T cell antigen-specific recognition of MHC class II restricted T cells; helper activity (stabilizes TCR-MHC class I interaction)</td>
<td>Ig SF 60 kDa transmembrane (TM) glycoprotein monomer</td>
<td>Saito et al., 1984</td>
</tr>
<tr>
<td>TCR α chain</td>
<td>Subtractive hybridization and Differential screening</td>
<td>Same subtracted library as Saito et al., above; differentially screened with two probes: subtracted probe generated in the same way as subtracted library and macrophase line-specific probe</td>
<td>T cell-restricted; individual components restricted to various thymocyte and T cell subsets</td>
<td>T cell antigen-specific recognition of MHC class II restricted T cells; helper activity (stabilizes TCR-MHC class I interaction)</td>
<td>Ig SF glycoprotein TM ab or as dimer of 34 kDa each</td>
<td>Saito et al., 1984</td>
</tr>
<tr>
<td>CD4</td>
<td>Subtractive hybridization</td>
<td>Subtracted probe made from mouse fibroblast L cells transformed with mouse T leukemic line Hut-102 (tester) subtracted with untransformed L cells-driven derived; used to screen a human PBL. T cell cDNA library</td>
<td>Pre-B cell specific</td>
<td>Required for early B cell development progression</td>
<td>Surrogate light chain component of pre-B cell receptor complex; homology to conventional F region</td>
<td>Chien et al., 1994</td>
</tr>
<tr>
<td>CD8</td>
<td>Subtractive hybridization</td>
<td>Subtracted probe made from mouse fibroblast L cells transformed with mouse T leukemic line Hut-102 (tester) subtracted with untransformed L cells-driven derived; used to screen a human PBL. T cell cDNA library</td>
<td>DP thymocytes, CTLs, some dendritic cells</td>
<td>TCR co-receptor molecule; signaling via p56lck; maturation and positive selection of MHC class II-restricted Th cells; helper activity (stabilizes TCR-MHC class I interaction)</td>
<td>Ig SF 60 kDa transmembrane (TM) glycoprotein monomer</td>
<td>Littman et al., 1985</td>
</tr>
<tr>
<td>A5</td>
<td>Subtractive hybridization and Differential screening</td>
<td>Subtracted library made from 70Z/3 mouse pre-B cell tester subtracted with Th cell hybridoma K62; differentially screened in duplicate with either 70Z/3-specific or K62-specific probes</td>
<td>Pre-B cell specific</td>
<td>Required for early B cell development progression</td>
<td>Surrogate light chain component of pre-B cell receptor complex; homology to conventional F region</td>
<td>Sakaguchi et al., 1988; [307]</td>
</tr>
<tr>
<td>CD22</td>
<td>Subtractive hybridization</td>
<td>Subtracted cDNA probes generated from P1A+PMA activated human tonsilular B cells (tester source) vs SAC+PM-activated human PBL. T cell (developer source); used to screen a normal SAC+PM-activated human tonsilular B cell cDNA library</td>
<td>B-cell specific; cytoplasmic expression in late pre-and early-pre-B cells; surface expression on resting mature B cells</td>
<td>Binds to a2-6 linked sialylated glycogenactan. Role in &quot;negative&quot; B cell signal: limiting BCR signal; also mediates adhesion to T cells, B cells, monocytes, and neutrophils</td>
<td>Ig SF, C-locus type I transmembrane glycoprotein with cytoplasmic ITIM negative signaling motifs</td>
<td>Wilson et al., 1991</td>
</tr>
<tr>
<td>LSP-1 (Lymphocyte-specific protein; aka pp52)</td>
<td>Subtractive hybridization</td>
<td>Subtracted library prepared from human CTL cell line AH2 tester subtracted with EBV- transformed B cell line I.B.1; screened with subtracted probe made in same way</td>
<td>Human and mouse homologues both restricted to lymphoid, macrophage, and neurophil lineages; predominant in lymphocyte tissues; in T, lineage restricted to DP thymocytes and SP T cells; also in B cell lymphomas and leukemias, but not in T cell lymphomas</td>
<td>T-cell binding protein associated with cytoskeleton and cytoplasmic side of plasma membrane; proposed role in mediating lymphocyte-specific cytokine-related structure/development; receptor interaction, cell motility, and cell-cell interactions; anti-tumorigenic role in T cell lymphomas?</td>
<td>No homology</td>
<td>Jongstra et al., 1987</td>
</tr>
<tr>
<td>GITR (glucocorticoid-induced TNP receptor family-related gene)</td>
<td>DD RT-PCR</td>
<td>Comparison of unreacted and dexamethasone-treated mouse T cell hybridoma (3DO) cells</td>
<td>T-cell restricted; increased expression with PMA+iono, antigen stimulation, or ConA stimulation</td>
<td>Inhibitor of TCR-mediated apoptosis</td>
<td>Tumor necrosis factor family type I transmembrane protein, high homology to CD27 and 41BB-Ligand</td>
<td>Nocentini et al., 1997 [327]</td>
</tr>
<tr>
<td>ABCD1</td>
<td>Supression</td>
<td>Subtraction of teter mRNA from mouse stromal cell-derived B cell line R2BFL-1.L-7 with tester mRNA from R2BFL-1.L-4+CD40-4.L-7</td>
<td>Lymph node tissue-restricted; preferentially activated splenic B cells and dendritic cells</td>
<td>Chemotaxis of ConA-2-activated splenic T cells</td>
<td>CC chemokine family: likely orthologue of MDCSTCP-1 human chemokine</td>
<td>Schaniel et al., 1998</td>
</tr>
<tr>
<td>FAIM (Fas Apoptosis Inhibitor Molecule)</td>
<td>DD RT-PCR</td>
<td>Comparison of mouse primary splenic B cells that are rendered Fas-resistant (via CD40L, stimulation) and those rendered Fas-sensitive (via CD40L,am-1gM stimulation)</td>
<td>Ubiquitous constitutively expression; inducible effector whose expression increases only in B cells</td>
<td>Specifically mediates Fas resistance in B cells produced by BCR ligation: decreased levels of PARP cleavage activity and cell death/lysis in response to Fas-L stimulation in FAIM-transfected cells</td>
<td>No homology</td>
<td>Schneider et al., 1999</td>
</tr>
<tr>
<td>Gene</td>
<td>Approach used for original isolation</td>
<td>General strategy/cell lines/tissues used for subtractions and/or comparisons in original screens</td>
<td>Reported expression pattern</td>
<td>Postulated/characterized function in lymphocyte development</td>
<td>General structural classification</td>
<td>Year identified/origin reference(s)</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>---------------------------------------------------------------</td>
<td>--------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Protein S</td>
<td>DD RT-PCR</td>
<td>Comparison of mouse primary T cells +/- anti-CD3+IL-2 (or IL-4) treatment</td>
<td>Constitutively expressed on hepatic tissue, testis, brain, ovary, thymus, and spleen; upregulated in IL-4 treated T cells</td>
<td>IL-4 inducible target; inhibition of lymphoid cell procoagulant activity</td>
<td>Coagulation pathway antagon agonist; proteolytically involved in inactivating coagulation factor prothrombinase</td>
<td>Smiley et al., 1997 [336]</td>
</tr>
<tr>
<td>SNAP-23</td>
<td>DD RT-PCR</td>
<td>Comparison of mouse IL-3 dependent OTt 1 transgenic myelomonocytic cell line +/- IL-3 treatment</td>
<td>Ubiquitous constitutive expression</td>
<td>Cytokine signal transduction downstream target of various immediate early cytokines including IL-2, 3, 5, 10, SCP, G-CSF; GM-CSF; and erythropoietin</td>
<td>87% homology to human SNAP-23</td>
<td>Morikawa et al., 1998</td>
</tr>
<tr>
<td>RsaN</td>
<td>DD RT-PCR</td>
<td>Comparison of human PHB+/- various T cell-activating agents</td>
<td>?</td>
<td>7 correlation in T cells with cell cycle decreased in G1/G2</td>
<td>Human retinoic acid receptor subunit</td>
<td>Isaza et al., 1998</td>
</tr>
<tr>
<td>CD36</td>
<td>Subtractive hybridization and Differential screening</td>
<td>Subtractive library prepared from tester Oct2-t Atelion pre-B mouse cell line which had inducible stable transfectant driver population; subsequent differential screening of positives with IL-3- or transfectant-specific cDNA probes</td>
<td>B cell lymphomas, myelomonocytic, platelets, megakaryocytes, and endothelial cells</td>
<td>Oct-2 downstream transcriptional target; Oct-2 mediated B cell differentiation via CD36 differentiative signals/co-stimulation and/or involved in membrane synthesis as fatty acid receptor</td>
<td>Membrane receptor implicated in mast cell-induced cytokine release, PTK associated signal transduction, and receptor and putative transporter for long chain fatty acid synthesis</td>
<td>Konig et al., 1995 [328]</td>
</tr>
<tr>
<td>GATA-3</td>
<td>cDNA-RDA</td>
<td>Naive mouse CD4 T cells differentiated in vitro into Th1 and Th2 populations with IL-12/- and IL-4, respectively; subtraction of IL-12 differentiated population</td>
<td>Hematopoietic cell-restricted, predominant in T and erythrocyte lineages; within CD4 T lineage, selectively expressed in Th1 cells</td>
<td>Required in Th2 cytokine transcription; as shown by antisense and transgene expression</td>
<td>Gata zinc-finger transcription factor family</td>
<td>Zheng et al., 1997 [335]</td>
</tr>
<tr>
<td>LGALS1</td>
<td>(lectin, galactoside-binding, soluble gene)</td>
<td>Comparison of mouse T cells +/- conA stimulation</td>
<td>High expression in activated T, but not in activated B, or resting B or T cells; general tissue distribution not reported</td>
<td>Autoimmune negative growth factor for CD4+ T cells</td>
<td>0-galactoside-binding protein family (GBGP) with lectin-like domains</td>
<td>Blaser et al., 1998</td>
</tr>
<tr>
<td>STZL</td>
<td>(aka T, DER4, Ftl1)</td>
<td>Balb/c mouse splenic naive CD4+ T cells differentiated in vitro into IL-12/- and IL-4/- (for Th1 line) or IL-4 (for Th2 line); comparison of Th1 and Th2 differentiated CD4 T cells</td>
<td>Expressed in hematopoietic lineages and predominantly restricted to adult bone marrow; inducible upon mitogen stimulation</td>
<td>Th2-specific signaling?</td>
<td>Ig superfamily member; high homology to type 1 IL-1 receptor</td>
<td>Xu et al., 1998 [334]</td>
</tr>
<tr>
<td>BSAP</td>
<td>(B cell-specific activator protein; aka Pax-5)</td>
<td>Comparison of low and high RAG-expressing variants of human mature B cell line OCI 1, 8, 3-SP</td>
<td>B cells, CNS, testis; co-expressed with RAG-1/2 in human B cell lines</td>
<td>Required for B cell differentiation; transcriptional regulation of B-specific target genes A5, VpreB, and CD19; possible additional function in B cell-specific RAG2 transcriptional regulation</td>
<td>B cell-specific paired domain Pax family transcription factor</td>
<td>Verkoczy et al., 1998 [392]</td>
</tr>
<tr>
<td>TCA-3 (T cell activating-3)</td>
<td>Subtractive hybridization</td>
<td>Subtracted probe generated from mitogen or antigen-induced mouse T cell hybridoma CI.11.Y-1 Ttager subfraction with mouse non-hematoid mixture of L cells, B cell lymphoma 2PKS, and B cell myeloma MOPC315, subtracted with unstimulated CI.11.Y-1 Ttager probe used to screen a ConA treated CI.11.Y-1 derived cDNA library</td>
<td>Restricted to ConA or anti-activated T cell clones</td>
<td>Novel T cell-activating lymphokine</td>
<td>69 as secreted polypeptide with no reported homology</td>
<td>Burd et al., 1987</td>
</tr>
<tr>
<td>YT15</td>
<td>Subtractive hybridization</td>
<td>Subtracted probes made from human T cell leukemia line Jurkat (tesser source) subfraction with driver derived from a mixture of 3 human B lymphoma cell lines; OCI 1 Y9, RPM11788, and RPM13638; used to screen various T and non-T cell line RNAs in Northern blots</td>
<td>Predominantly thymocyte-restricted</td>
<td>T cell activation</td>
<td>No homology</td>
<td>Takahara et al., 1989</td>
</tr>
<tr>
<td>LAG-1 (lymphocyte activation Gene 1)</td>
<td>Subtractive Hybridization</td>
<td>Subtracted cDNA probe made from human F5 NK-cell line tester subfraction with human Jurkat (T) K562 (erythroid-myeloid) cell line; used to screen a F5-derived cDNA library</td>
<td>General expression pattern not reported; constitutive expression in all T cells, increased levels in stimulated T and NK cells</td>
<td>T cell activation</td>
<td>69 as secreted polypeptide belonging to RANTES lymphokine family</td>
<td>Baixeras et al., 1990</td>
</tr>
<tr>
<td>PD1</td>
<td>Subtractive hybridization and Differential screening</td>
<td>Subtracted probe made from apoptosis induced (via IL-3 withdrawal) IL-3-dependent lymphoblastoid progenitor mouse bone cell line LYN9 subtracted with LYN9 supplemented with IL-3 driver population; used to screen subtracted library created from PMA, tumor necrosis factor induced-activated T cell line with untreated 240B.11 mouse T cell hybridoma driver population; clones isolated differentially screened with LYN9 subtracted probes or LYN9-specific probes</td>
<td>Predominantly DN thymocyte-restricted</td>
<td>Thymocyte apoptosis</td>
<td>Novel Ig superfamily member</td>
<td>Ishida et al., 1992</td>
</tr>
<tr>
<td>BL11</td>
<td>Subtractive hybridization</td>
<td>Screening of CD28+ SAC+PMA treated human tonsillar B cell cDNA library with subtracted probe prepared from CD28+ SAC+PMA treated human tonsillar B cells (tesser source) vs. PHA+PMA-treated human CD3+ T cells (driver source)</td>
<td>Preferentially expressed in mitogen-activated B cells and lung</td>
<td>B cell activation</td>
<td>Ig superfamily glycoprotein</td>
<td>Kozlow et al., 1993</td>
</tr>
<tr>
<td>BL34</td>
<td>Subtractive hybridization</td>
<td>Preferentially expressed in mitogen-activated B cells; upregulated in A.L. patients</td>
<td>No homology</td>
<td>No homology</td>
<td>No homology</td>
<td>Hong et al., 1996</td>
</tr>
<tr>
<td>Gene</td>
<td>Approach used for original isolation</td>
<td>General strategy/cell line/ tissues used for subtractions and/or comparisons in original screens</td>
<td>Reported expression pattern</td>
<td>Postulated/characterized function in lymphocyte development</td>
<td>General in structural classification</td>
<td>Year identified/ original reference(s)</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td>----------------------------</td>
<td>------------------------------------------------------------</td>
<td>------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>CTLA-8</td>
<td>Subtractive hybridization and Differential screening</td>
<td>Screening of subtracted library duplicates (library generated from subtraction of PMA+ionomycin treated d10s mouse CTL hybridoma clones toest reverse hybridization) with various combinations of subtracted probes generated from same mouse lymphoma cells used to make libraries</td>
<td>Restricted exclusively to PMA+ionomycin treated T cell lymphomas d10s and d11s, used to prepare libraries</td>
<td>T cell death/survival?</td>
<td>57% homology to Saimiri herpesvirus ORF13 gene</td>
<td>Rouvier et al., 1993</td>
</tr>
<tr>
<td>Ly-GDI</td>
<td>Subtractive hybridization</td>
<td>Screening of human BJAB B lymphocyte cell line library with subtracted probes prepared from human K562 myeloid cell line-deleted teater vs. Raji/EW B lymphoma cell line-derived driver</td>
<td>Hematopoietic tissue- restricted; predominately in B cell line/tissues</td>
<td>B221 cell activation?</td>
<td>78% homology to Rh-GDI, an inhibitor of GTPase exchange in Rh GTPases</td>
<td>Scherle et al., 1993</td>
</tr>
<tr>
<td>LSM1</td>
<td>Subtractive hybridization</td>
<td>cDNA libraries constructed from mouse pre-B Abelson transformed temperature sensitive mouse pre-B cell line subclone SPL-212-12+ (temperature shift driver/+)-tector</td>
<td>Restricted to lymphocyte mouse adult cell lines and spleen, thymus, BM, embryo tissues</td>
<td>?</td>
<td>transmembrane protein with no homology</td>
<td>Shimizu et al., 1994</td>
</tr>
<tr>
<td>M17</td>
<td>Subtractive Hybridization</td>
<td>Subtraction of PNA mouse GC B cell tester population with PNA mouse GC B cell driver; subtraction of PCR-amplified cDNA libraries derived from single cells</td>
<td>GC mature B cell- specific</td>
<td>?; signal transduction</td>
<td>cytoplasmic protein with no homology; has lipid-binding and ITAM motifs</td>
<td>Christoph et al., 1994</td>
</tr>
<tr>
<td>CRISP-3</td>
<td>Subtractive hybridization and Differential screening</td>
<td>Same as for CD36</td>
<td>Restricted to low in spleen and high levels in bone marrow and spleen, pre-B mouse cell line-specific, no expression in B, plasma, non-lymphoid or T cell lines</td>
<td>?; downstream transcriptional target of Oct2; anti-fungal/antibacterial functions in blood or lymph?</td>
<td>novel member of CRISP subfamily; high homology to plant defense protein family</td>
<td>Pfisterer et al., 2001 [329]</td>
</tr>
<tr>
<td>B4B</td>
<td>DD RT-PCR</td>
<td>Comparison of low, intermediate, and high density fractions of human PBMC and EBV-transformed B cell line REM</td>
<td>Expressed at variable levels in most hematopoietic cell lines, but within B lineage, restricted to cytoplasmic pre-B cell subset</td>
<td>Growth arrest in COS-7 transfected cells; elimination of B cells before heavy chain rearrangement and growth arrest of transformed progenitor B cells</td>
<td>tetra-span transmembrane protein</td>
<td>Ruegg et al., 1996 [332]</td>
</tr>
<tr>
<td>C2.3</td>
<td>DD RT-PCR</td>
<td>Comparison of PA6 mouse RAG-inducing stromal cell line PA6 RAG(+) - RAG(-) non-inducing variant</td>
<td>Mouse spleen, thymus, testis, kidney, lung, heart, bone</td>
<td>B221 cell death 7; stromal cell surface receptor that may deliver signal for RAG induction in pre-B cells</td>
<td>tetra-span transmembrane protein</td>
<td>Tagoh et al., 1996 [216]</td>
</tr>
<tr>
<td>Fig1</td>
<td>(interleukin-four induced gene 1) cDNA-RDA</td>
<td>Subtraction of LPS/+/-IL-4 (driver/+)-tector stimulated mouse splenic B cells</td>
<td>Lymph node tissue-restricted; inducible by IL-4 only in B cells</td>
<td>?; immediate-early IL-4 inducible gene</td>
<td>Novell protein with homology to monamine oxidases; high homology in PAD binding domains</td>
<td>Chu and Paul, 1997 [326]</td>
</tr>
<tr>
<td>RPI</td>
<td>DD RT-PCR</td>
<td>Comparison of human resting PMBC T cell populations with T cells stimulated by anti-CD3 crosslinking, anti-CD28 crosslinking or both</td>
<td>Activated T cells, general expression pattern not reported</td>
<td>?; TCR signal transduction</td>
<td>indeterminate polyclonal cell (APC) protein-binding EBI family</td>
<td>Renner et al., 1997</td>
</tr>
<tr>
<td>HDI (histone desaturase gene)</td>
<td>DD RT-PCR</td>
<td>Comparison of mouse T cells +/- IL-2 treatment</td>
<td>General expression pattern not assessed, high expression in activated T cells at G1/S cell cycle boundary</td>
<td>?; histone deacetylase activity, cell cycle regulation in lymphocytes?</td>
<td>nuclear protein with high homology to yeast RPD3 pleiotropic transcriptional regulator</td>
<td>Bartl et al., 1997</td>
</tr>
<tr>
<td>Decysin</td>
<td>SSH</td>
<td>Subtraction of CD40-activated human tonsilar GC Dendritic cells, high expression with U937 human histiocytic cell line- derived driver population</td>
<td>Upregulated only in CD40-activated GC dendritic cells, high expression with U937 human histiocytic cell line-derived driver population</td>
<td>Dendritic cell T cell interactions- selection/ activation; implicated in TNF-α regulation and Fas-1 processing</td>
<td>novel disintegrin metalloproteinase family member</td>
<td>Mueller et al., 1997 [333]</td>
</tr>
<tr>
<td>hBRAG (human B cell RAG-associated gene)</td>
<td>DD RT-PCR</td>
<td>Same as for BSAP</td>
<td>High expression in spleen, LN, low expression in other tissues.expressed with RAG in human B cell lines</td>
<td>B cell-specific RAG regulation?</td>
<td>type II transmembrane glycoprotein</td>
<td>Vorkoczky et al., 1998 [45]</td>
</tr>
<tr>
<td>TOAM7 (T cell growth associated molecule 77)</td>
<td>DD RT-PCR</td>
<td>Comparison of resting PB human T cells to those in vitro activated at different time points</td>
<td>Variable expression in numerous tissues</td>
<td>Localization of T cell activation signaling events?</td>
<td>65% homology to D. Melanogaster cornichon protein</td>
<td>Usho et al., 1999</td>
</tr>
<tr>
<td>ST6GalNacL (T cell growth associated molecule 77)</td>
<td>DD RT-PCR</td>
<td>Comparison of mouse C57 T cells +/- nitro- gen or anti-N-ger stimulation</td>
<td>Ubiquitous expression</td>
<td>Alteration of sialylation pattern of cell surface molecules in activated lymphocytes?</td>
<td>c2-o sialyltransferase</td>
<td>Kaufmann et al., 1999</td>
</tr>
<tr>
<td>BLR1 (Burkitt's lymphoma receptor 1)</td>
<td>Subtractive hybridization</td>
<td>Cyclohexamide-treated human Burkitt's lymphoma (BL) B cell line BL64 vs. human primary BL tumor cells</td>
<td>Exclusively expressed in BL cell lines and, lymphomas</td>
<td>Mature B cell activation? cell-cell interactions, tissue-specific?</td>
<td>novel G protein family member; high homology to IL-8 and other neutrophil chemotaxaotrans with high oncogenic oncogenic potential</td>
<td>Dobner et al., 1992</td>
</tr>
<tr>
<td>ARF (All responsive gene 1; aka RIEG, Prx2, Otx)</td>
<td>DD RT-PCR</td>
<td>Comparison of mouse ALL1.1-/- vs. WT embryonic stem cells</td>
<td>Predominantly in normal human bone marrow; eye, brain, and stomach; brain tissues; not expressed in ALL lines with 11q23 translocations</td>
<td>?; target of ALL1, the human homologue of Drosophila trithorax gene, involved in ALL abnormalities at 11q23</td>
<td>high homology to Ptx1/Ptx2 homoeic gene</td>
<td>Arakawa et al., 1998</td>
</tr>
</tbody>
</table>

1. categories of genes listed in order of year isolation  
2. whether not reported or postulated/speculated function

Table 1. cont'd
novel genes specifically expressed in response to cytokine induction. For example, in B cells, *Fig1* has been identified as the first immediate-early gene induced by IL-4 [326]. Similarly, in T cells, a new TNF family member, GITR, has been identified as a dexamethasone-inducible gene [327]. With respect to the identification of transcription factor target genes (both novel and known), a successful approach has been to use DD RT-PCR or cDNA-RDA to assess altered gene expression from knockout cells. For example, gene expression analysis in Oct2⁺ vs. WT mice by differential screening identified CD36, and a novel gene, CRISP-3, to be downstream targets [328, 329]. Similarly, several novel candidate B cell-restricted targets of the B/myeloid-lineage transcription factor PU.1 have been identified using cDNA-RDA analysis of PU.1⁺ mice [330]. Future studies of this sort using microarray or SAGE analysis may reveal a comprehensive picture of all transcription factor target genes. One study has already used microarray to identify both known genes and unknown ESTs as targets of heat shock induction or PMA stimulation in T cells [331].

Another effective use for DD RT-PCR or other differential expression screening techniques has been in identifying novel B or T cell-specific genes via comparison/subtraction of T and B lineages, or discrete developmental stages/subsets within these lineages. For example, the pre-B cell restricted SLC component λ5 was identified over ten years ago by screening a 70Z/3 pre-B cell-specific subtracted library [97] while more recently, a novel pre-B cell specific growth arrest gene, B4B, has been identified by DD RT-PCR of different PBMC density fractions [332]. Other examples include the identification by cDNA-RDA of decysin, a novel disintegrin metalloproteinase expressed specifically in GC dendritic cells [333], and identification of ST2L, a molecule preferentially expressed in activated T₁₂ cells and associated with T₁₂ functions, isolated by DD RT-PCR comparison of activated T₁₁ and T₁₂ CD4⁺ T cell subsets [334]. Another related application has been to compare/subtract a spontaneously-arising T or B cell variant that has altered mRNA expression of a lymphocyte-restricted gene/marker because of altered gene expression in an upstream regulator. One interesting example using DD RT-PCR for this purpose is a study in which mRNA expression of a potential regulator of RAG in early, stromal-cell dependent pro-B cell development is lost [217]. In this system, a combination of cytokines in conjunction with stromal cells induces expression in mouse pro-B cells (see section 1.2.1.1). A spontaneously-arising variant incapable of mediating RAG1 induction was
isolated and postulated to lack expression of a cell surface receptor required for this induction. By DD RT-PCR comparison of the inducible parental line and the non-inducible variant, a cDNA encoding a novel tetra-spanning transmembrane molecule, C2-3, was isolated and found capable of restoring RAG1 induction when transfected back into the RAG non-inducing variant.

DD RT-PCR and other differential expression analysis techniques have not only been useful for isolating new genes, but also for identifying new roles or associations for known genes (upstream or downstream of known or novel genes). For example, by using cDNA-RDA to compare T H2 vs. T H1 subsets, the zinc-finger transcription factor GATA-3 was shown to be selectively expressed in T H2 cells, suggesting its preferential involvement in this subset [335]. Several follow-up experiments, including GATA-3 antisense and transgenic mouse experiments, all subsequently confirmed that GATA-3 was shown to be necessary and sufficient for T H2 type cytokine gene expression in CD4+ T cells. Similarly, DD RT-PCR revealed that Protein S, a known physiological anticoagulant, is inducible by IL-4 in T cells [336]. Additionally, though DD RT-PCR comparison of IL-2 and IL-12-induced T cells, it has been shown that IL-12 activation, unlike IL-2 induction, does not involve immediate-early gene expression. This suggests that differential gene induction may result in the functional differences between the partially overlapping IL-2 and IL-12 responses [337].

Finally, DD RT-PCR analysis has also shed potential molecular clues on the pathogenic mechanisms of lymphocyte disease or infection. For example, in one recent study, DD RT-PCR revealed that EBV transformation of B cells will downregulate TAP1 expression (transporter associated with antigen presentation) [338]. The newly discovered downregulation of this gene, encoding a protein involved in the MHC antigen-processing and presentation pathway, may help to explain why EBV lowers viral presentation.
1.4 Type II transmembrane proteins with characterized immune functions

Because a large portion of my results will deal with hBRAG, one of genes I identified by the above-described DD RT-PCR methodology, and it encodes a novel protein with a type II TM, other type II molecules with immunologic functions will be briefly reviewed. Overall, type II proteins are relatively rare as they comprise only ~5% of TM proteins overall (reviewed in [339]). There are basically two major families of type II TM molecules with common structure and similar immune functions: the tumor necrosis factor (TNF) ligand family and the C-lectin domain type II TM multigene family. Three other unique type II TM proteins belonging to completely distinct classes structurally and functionally are: the BP-1 antigen, the MARCO/macrohage scavenger type I receptors and the MHC class II-associated invariant chain (Ia). Figure 1.17 shows the structural classification of these various type II TM proteins.

1.4.1 The TNF ligand family

All TNF ligand subfamily members are type II TM proteins expressed in the T cell lineage (although not uniquely so), and almost all have co-stimulatory functions [340]. This family is part of the larger TNF/Nerve Growth Receptor (NGF) superfamily which includes the corresponding receptors of these ligands [341]. All TNF superfamily members are cysteine-rich molecules which share homology primarily in the C-terminus (the extracellular receptor-binding portion), ultimately resulting in a TNF-specific β sandwich tertiary structure. It is becoming clear that referring to TNF ligand subfamily members as ligands is a misnomer since many of these products appear to have signaling potential, and as such, may be more appropriately termed counter-receptors. Additionally, some of these ligands also exist in alternative soluble forms (good examples are TNFα and β) and as such can serve as regulators of the corresponding ligand isoform. Some TNF ligand family members are involved in T-T interactions, such as the CD27L (CD70) which enhances T cell proliferation and cytolytic T cell activity, while others are involved in T-B interactions, such as CD40L, which provides T cell help to B cells. Other members include CD30L, CD134L (OX40L), 41-BBL, and the TNFα and β subcomponents (Figure 1.17). The best-characterized amongst these co-stimulatory products is the CD40L (also known as CD154 or gp39), whose corresponding mutations have been functionally
<table>
<thead>
<tr>
<th>TNF ligand family</th>
<th>C-lectin receptor family</th>
<th>Scavenger receptor family</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homotrimer</td>
<td>Homotrimer</td>
<td>Homotrimer</td>
</tr>
<tr>
<td>NH$_2$</td>
<td>H$_2$N NH$_2$</td>
<td>NH$_2$NH$_2$NH$_2$</td>
</tr>
<tr>
<td>TNF</td>
<td>LY49-A</td>
<td>NKG2</td>
</tr>
<tr>
<td>FASL</td>
<td>CD69</td>
<td>CD23</td>
</tr>
<tr>
<td>4-IBBL</td>
<td>CD72</td>
<td>Macrophage Lectin</td>
</tr>
<tr>
<td>CD27L</td>
<td>CD94</td>
<td></td>
</tr>
<tr>
<td>CD30L</td>
<td>CD161</td>
<td></td>
</tr>
<tr>
<td>CD40L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD134L (OX40L)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1.17. Structures of various type II transmembrane protein family members with immunological function and/or expression. The three major families are shown. A legend representing the various family-specific domains is shown at the bottom. Not shown are the BP-1 molecule, an ecto-metallopeptidase of unknown function expressed in early B cell development, and invariant chain (Ia), an intracellular type II protein involved in MHC class II transport/assembly.
linked to the human immunodeficiency called hyperIgM syndrome (HIM) [27-29]. Another intensively characterized molecule with a very different function from the other members of the TNF ligand family is Fas-L, (also called CD95L). This protein has an important role in the delivery of a potent apoptosis signal via Fas ligation (reviewed in [342]). Structurally, it can be distinguished from other members by its death domain. A naturally occurring Fas-L deficient mouse, the gld mutant, shows an autoimmune phenotype, similar to the Fas-deficient mutant mouse lpr.

1.4.2 The LY49 family

The second major family of type II TM proteins, the C-lectin type II TM family (including members of the LY49 multigene receptor family with inverted orientation), is characterized by having a domain in their C-terminal extracellular regions shared by all C-lectin superfamily members, the C (carbohydrate)-lectin recognition domain (CRD) (reviewed in [343]). CRD regions are the most highly conserved regions of these members, but there are also somewhat lower homologies which exist in the cytoplasmic domain and extracellular domain proximal to the TM region (~80% overall aa identity). The CRD regions are believed to be involved in-carbohydrate binding but also have other functions; for example, the CD23 CRD is thought to bind IgE Fc [344]. C-lectin superfamily members are generally large, heavily glycosylated polypeptides with multiple CRD domains. This is in contrast to the C-lectin type II TM family within this superfamily, which unlike the other superfamily members, are generally smaller polypeptides with only one CRD domain and are generally not heavily glycosylated (containing usually only one N-glycosylation site). Many of these, including the mouse NK and T cell subset-specific genes LY49, CD94, NKG2 and NKR-P1, are also part of another emerging family of receptors, the NK inhibitory receptor family which include many recently cloned molecules including the human KIR (Killer Inhibitory Receptor), PIR (Paired Ig Receptor), MIR (Monocyte Ig-like Receptor) and LAIR (Leukocyte-Associated Ig-like Receptor) subfamilies (reviewed in [345]). Many of these molecules, including LY49, CD94, are ITIM-containing signal transmitting receptors involved in negative signaling instead of co-stimulatory or "positive" signaling (reviewed in [128]). Other C-lectin members include a diverse array of molecules: the low-affinity IgE receptor (CD23), the IL-1R α and β polypeptide chains, and CD69 (which has co-stimulatory function and is
expressed preferentially on activated lymphocytes and NK cells). They also include the B-cell specific signaling receptors, CD23 (which is expressed preferentially on activated B cells and monocytes), and the CD23-homologous B cell differentiation antigen CD72 (Lyb-2). CD23<sup>−</sup> mice generally exhibit normal immune responses lymphocyte differentiation, but have elevated and sustained levels of IgE antibody titers in response to T-dependent antigens [346]. Similar to the KIR and LY-49 members, and the earlier-described CD22, these two proteins are involved in negative signaling via their respective ITIMs and subsequent recruitment of the shp-1 phosphatase [347], which negatively regulates ITAM activation by dephosphorylating signaling intermediates (reviewed in [128]).

1.4.3 Miscellaneous other type II TM proteins with characterized immune functions

**BP-1 antigen**

The BP-1 antigen (also called glutamyl endopeptidase/6C3/EAP) is a homodimeric disulfide-linked protein composed of two 140 kDa type II polypeptide subunits which are members of the aminopeptidase member group within the metalloproteinase family [348]. As described earlier, BP-1 has a similar structure and expression pattern as the human endopeptidase CD10 and within the B lineage, is an important marker of early mouse B cell differentiation, restricted to pro-B/pre-B murine B cell subsets (for expression pattern see Figure 1.3), and can be upregulated by IL-7 or viral transformation. As an aminopeptidase, BP-1 is an ecto-enzyme (cell-surface expressed) and has a zinc-binding motif shared by other aminopeptidases that is required for their enzymatic activity. Aminopeptidases work by cleaving acidic aa residues from polypeptide substrates. The role of BP-1 in B cell development still remains elusive but several mechanisms by which it may act have been proposed (reviewed in [139]). BP-1 has been postulated to have a role in cleaving stromal cell regulatory peptide factors from their surface that act either on pre-B or stromal cells; in addition it may modify the IL7/IL7R interaction since the IL-7R has an N-terminal glutamic acid residue which is substrate for BP-1. Alternatively, BP-1 may also be an adhesion molecule for BM pre-B cells and have an enzyme-independent function. BP-1<sup>−</sup> mice display normal development and lymphocyte responses, and thus may be compensated for by another peptidase ecto-enzyme [349].
Scavenger receptor type I

Macrophage scavenger receptors type I and II (called MSR-AI and MSR-AII, respectively), are part of a growing family of receptors with SRR (scavenger receptor cysteine-rich) domains, first recognized by Freeman et al. ([350]). MSR-AI and MSR-AII are restricted to cells of monocyte origin and are most highly expressed during the process of monocyte to macrophage differentiation (reviewed in [351]). Of these, those with type II TM structure belong to the MSR-AI group, which differs from the later-discovered MSR-AII group only in the presence its C-terminal SRCR domain, which has a cysteine-rich region [352]. These include scavenger R type I (MSR-A1) [353] and the macrophage receptor with collagenous structure (MARCO) [354, 355]. MARCO and MSR-A1 are trimeric type II membrane glycoproteins containing three C-terminal extracellular cysteine-rich domains connected to the TM domain by a unique long, fibrous, heavily-glycosylated stalk structure, composed of an alpha-helical coiled coil and a collagen-like triple helix (Figure 1.17). In contrast to MARCO and MSR-A1, other SCRC family members are type I TM glycoproteins, and instead of having a single C-terminal SRCR, have extracellular domains entirely composed of multiple SRCR domains. For example, CD6 and CD163, the former involved in T cell activation, have three and nine SCRC domains, respectively.

These receptors have extraordinarily promiscuous binding specificity. The list includes polyanions, low-density lipoproteins, and lipoprotein-rich bacterial pathogens (for e.g., lipoteichoic acid of gram-positive bacteria). The binding of the latter suggests the mechanism of MSR-A1-mediated host defense. In this context, mice with a targeted disruption in MSR-A1 infected artificially with BCG are more susceptible to endotoxic shock and produce more TNF-α and IL-6 [356]. MSR-A1⁺ mice are also more naturally susceptible to Listeria and HSV type I [357]. Mechanistically, it is likely that the MSR-A1 SRR plays a protective role by directly blocking or lowering release of pro-inflammatory cytokines, and/or by sequestering or "scavenging" LPS, which reduces LPS-mediated toxicity via lowering monocyte overactivation. Both MSR-AI and MSR-A2 are mediators for the binding and internalization of chemically-modified (oxidized, low-density) lipoproteins. Atherosclerosis patient lesions accumulate such proteins, implying a role for defective scavenger receptor expression and/or function in atherosclerosis pathogenesis [357]. Supporting this are studies in Lipoprotein A+MSR-A1⁺ mice, which have significant reductions in atherosclerotic lesions relative to Lipoprotein⁺ mice [357].
Invariant chain

The only known example of a predominantly intracellularly expressed type II TM protein with a characterized immune function is the invariant chain (Ia), involved in MHC class II transport/assembly (reviewed in [358]). In the overall class II assembly and intracellular transport pathway, Ia-class II complexes are generally believed to exit the ER, travel through the Golgi, and are targeted to endosomes. Ia is then proteolytically cleaved, releasing class II dimers. Ia is co-regulated with MHC class II genes and is expressed in all class II positive cells i.e. B cells and APC-macrophages and dendritic cells. The Ia polypeptide is a unique type II TM protein that has a short (~30 aa cytoplasmic domain, and a large, ~160 aa lumenal carboxyl-terminal region with two N-glycosylation sites). The endoplasmic reticulum (ER) retention signal and the endosomal targeting signals are both located in the N-terminal cytoplasmic region. When the endosomal-targeting region is removed, Ia is predominantly cell surface expressed, rather than intracellular. The main isoform, p33 is 33 kDa, but alternative isoforms exist including p35, p41, and p43, the molecular weight heterogeneity due to alternative translation initiation sites, alternative splicing, multiple proteolytic cleavage sites, and glycosylation differences. One larger human-specific isoform, p41, can be expressed on the cell surface and may have an additional adhesion function in CD44 binding via its additional chondroitin sulfate residues. Normally, the Ia protein, in the absence of class II association, exists as a homo-trimeric core. After synthesis in the ER, Ia combines with the MHCII heterodimer to form a nonameric complex such that the invariant chain trimer is believed to form the core, surrounded by the three associated αβ dimers.

A number of functions for Ia in the MHCII biosynthesis/transport pathway have been proposed (reviewed in [359]). One function of Ia is that of a chaperone molecule in the proper assembly and conformational folding of MHC class II, as shown experimentally in mutant cells deficient in Ia. Another is in the intracellular trafficking/transport of class II molecules including egress from the ER, movement through the Golgi, and targeting to the late endosomes. Ia may also act as a competitive inhibitor by blocking MHC class II peptide sites until classII subunits enter endosome compartments as well as by targeting classII subunits to endosomal compartments in order to prevent premature peptide binding in the ER. However, Ia is not absolutely required for MHCII surface expression and function, since Ia knockout mice show reduced, but not absent levels of MHC class II surface expression and can present antigens, albeit less efficiently [360].
Thesis Summary

In Chapter 2, I have exploited low and high RAG-expressing human mature B cell line variants in a molecular approach to isolate genes differentially expressed along with RAG1 mRNAs. Using differential display reverse transcriptase PCR (DD RTR-PCR) methodology, I report the isolation of six candidate genes that are associated (either directly or inversely) with RAG expression and early B cell development. Interestingly, one of these genes was RAG2, but in addition I detected BSAP, a potential late B cell-specific regulator of RAG expression, and four other novel cDNAs. I also report my limitations with the differential display technique and modifications in further screening and purification steps I have undertaken subsequent to differential display itself in order to obtain single, uncontaminated bona fide cDNAs.

In Chapter 3, I further characterize one of the differential display cDNA products isolated from my screens in Chapter 2. This leads to the molecular cloning of a novel gene named hBRAG (human B cell RAG-Associated Gene). In characterizing this gene, I show that hBRAG is co-expressed with RAG1 in human B cell lines, is found at highest levels in B cell enriched tissues, is conserved across vertebrate species, and has a role in the positive regulation of RAG1 mRNA expression, possibly in a B-cell specific context. The B cell-specific activation of RAG by hBRAG and its preferential expression in peripheral B cell-enriched tissues suggests that it, like BSAP, may be involved in the late B cell-specific regulation of RAG expression.

In Chapter 4, I present data reporting biochemical characterization of the hBRAG protein using purified polyclonal antibodies I have generated against N and C-terminal regions of the hBRAG peptide. The binding of these antibodies are tested in various contexts, and these reagents are subsequently used in peptide blocking experiments to assess hBRAG protein expression pattern and cellular distribution. The glycosylation status of hBRAG is also assessed using these antibodies. Taken together, these studies suggest that hBRAG is a membrane-integrated surface-expressed glycoprotein dimer, consistent with earlier structural predictions of 4 N-glycosylation sites in the hBRAG extracellular domain. Furthermore, by co-immunoprecipitation, I also show that other proteins associated with hBRAG may be phosphorylated and possibly de-phosphorylated upon BCR
ligation. These may be novel proteins, as co-immunoprecipitations with BCR-associated kinases including Hck, Lyn, and Fyn, were found not to associate significantly with hBRAG. Preliminary results also suggest that hBRAG is capable of generating transducing signals either alone, and possibly in conjunction with the BCR. Based on these data, I hypothesize that hBRAG is a cell-surface receptor molecule that mediates potentially important regulatory signals in B cell development and/or B cell-specific RAG expression.

Other ways to characterize hBRAG function and those of the other isolated cDNAs are proposed in the General Discussion (Chapter 5). The overall implications of hBRAG, BSAP, and other potential factors in the regulation of RAG transcription as well as their putative roles in B cell leukemias, immunodeficiencies, and autoimmune conditions are also discussed. Finally, my experience with DD RT-PCR is discussed in the context of potential problematic sources and overall refinements to the technique.
CHAPTER 2:

Isolation of genes negatively or positively co-expressed with human Recombination Activating Gene 1 (RAG1) by Differential Display PCR (DD RT-PCR)
Acknowledgements

Much of the material in this chapter appears in the article entitled *Isolation of genes negatively or positively co-expressed with human Recombination Activating Gene 1 (RAG1) by Differential Display PCR (DD RT-PCR)* in Nucleic Acids Research, volume 26 (19), pages 4497-4507, by Laurent K. Verkoczy and Neil L. Berinstein (1998). I performed all experiments described in this chapter. The ACGT Corporation did sequencing of differential display cDNAs. This work was supported by the National Cancer Institute of Canada (NCIC grant # 7286).
2.1 ABSTRACT

Differential display (DD RT-PCR) has been extensively used for analysis of differential gene expression, but continues to be hampered by technical limitations that impair its effectiveness. In order to isolate novel genes co-expressing with human RAG1, I have developed an effective, multi-tiered screening/purification approach which effectively complements standard DD RT-PCR methodology. In "primary" screens, standard DD RT-PCR was used, detecting 22 reproducible differentially expressed amplicates between clonally related cell variants with differential constitutive expression of RAG mRNAs. "Secondary" screens used differential display (DD) amplicates as probes in low and high-stringency Northern blotting. 8/22 independent DD amplicates detected nine independent differentially expressed transcripts. "Tertiary" screens used reconfirmed amplicates as probes in Northern analysis of multiple RAG- and RAG+ sources. Reconfirmed DD amplicates detected six independent RAG co-expressing transcripts. All DD amplicaters reconfirmed by Northern blot were a heterogeneous mixture of amplicons, necessitating further purification for isolating single amplicons prior to subcloning and sequencing. To effectively select the appropriate amplicons from DD amplicates, I excised and eluted the amplicons directly from regions of Northern blots in which differentially expressed transcripts had been detected. Sequences of six purified amplicons specifically detecting RAG1 co-expressing transcripts included matches to portions of the human RAG2 and BSAP regions and to four novel partial cDNAs (three with homologies to human ESTs). Overall, my results confirm the power of DD RT-PCR but also suggest that even when using clonally related variants from the same cell line in addition to all appropriate internal controls previously reported, further screening and purification steps are still required in order to efficiently and specifically isolate differentially-expressed genes by DD RT-PCR.
2.2 INTRODUCTION

Central to V(D)J recombination are the products of the recombination Activating Gene1 (RAG1) and recombination Activating Gene2 (RAG2) locus [311, 361]. Their necessary role in this process has been well established based on several independent lines of evidence [38, 50, 83, 158, 186, 362]. The expression of RAG1 and RAG2 is precisely regulated. For the most part, RAG1 and RAG2 are only expressed concordantly and stage-specifically in cells of the lymphocyte lineage [37]. Briefly, a first "wave" of RAG expression is detected in committed B and T progenitor lymphocytes at the point where immunoglobulin (Ig) μ and T Cell Receptor (TCR) β and δ loci undergo V(D)J rearrangement [39, 53]. The second "wave" corresponds to V(J) rearrangement of the TCR-α [41] and Igκ and λ loci [42, 52]. A third "wave" of RAG expression has been observed by several groups in Ig+ B cells [62, 65, 71, 72, 75, 223] and TCR+ thymocytes [222, 224]. Recently, functional recombinase activity has been demonstrated to accompany a fourth "wave" of RAG expression in germinal centre mature B cells [248, 249].

The importance of precise regulation of RAG expression is obvious in cases where there is too little or no expression. For example, the resulting phenotypes in humans or mice in which the RAGs have been disrupted by mutation or by targeted recombination, respectively, is a lack of mature lymphocytes and a resulting severe combined immunodeficiency [50, 83, 363, 364]. In contrast, improper shutting off of the RAGs may potentially lead to aberrant rearrangements and subsequent oncogenic events, since the overexpression of the RAG locus in transgenic mice results in various lymphocytic abnormalities [58, 59].

Genes co-expressed with RAG may be of considerable general interest because they may represent unidentified, lymphocyte-specific components of the recombinase machinery and/or novel, developmentally regulated, lymphocyte-specific genes, some of which may themselves be regulating RAG mRNAs. Our overall objective was to isolate such genes. In this report, I have used the stable, constitutive differential expression in OCI LY8 to isolate by differential display
(DD RT-PCR) several partial cDNAs that detect RAG1 co-expressed transcripts. I report our limitations with the DD RT-PCR technique, and demonstrate modifications that have enabled us to identify six partial cDNAs that are co-expressed with RAG1 mRNA. Based on the sequences of these cDNAs and the expression pattern of transcripts that they detect, I suggest that the screening approach reported here may yield the isolation of other genes that are developmentally- and lineage-restricted in the same manner as the RAG genes.
2.3 MATERIALS AND METHODS

Cell lines, tissue culture, and RNA isolation

OCI LY8-C3P (μλ+) is a single-cell clone from the OCI LY8 human mature B cell line (IgM+/IgD- /CD10+/CD19+/CD20+/CD38+), a line originally established from an adult patient with a diffuse large cell lymphoma [365]. The OCI LY8 system has been detailed elsewhere [223]. RAG+ variants clonally derived from OCI LY8-C3P included C3-A1N and A8-6P. Other human B-lineage cell lines used included the diffuse large cell lymphoma lines OCI LYL and OCI LY2, the Burkitt's lymphoma lines Raji and Daudi and the human pre-B lines Nal-1 and PB-697 (pre-B acute lymphocytic leukemia, RAG1+, λ5+, μ+). Human T cell and non-lymphoid cell lines used in our studies include Jurkat, K562 (erythroleukemia, ATCC CCL 243), HeLa (ATCC, epithelial carcinoma) and U937 (pro-monocytic). All cells were routinely cultivated at 37°C/5% CO₂ in RPMI 1640 medium (ICN, St-Laurent, QUE.), supplemented with 10% fetal bovine serum (HyClone, Logan, UT), 2mM L-glutamine, 100 U/ml penicillin, and 100 μg/ml streptomycin (Gibco BRL, Grand Island, NY). For all cell lines, total RNA was extracted from cells in log-phase growth (5x10⁵ cells/ml) by the single-step guanidium thiocyanate phenol-chloroform extraction procedure [366].

DD RT-PCR

DD RT-PCR was performed on DNA-free total RNA from OCI LY8-C3P and C3-A11N (as described in section 1.3.4.1 of the general introduction and [309, 311]) with the following specific parameters. 40 primer combinations were employed, representing statistical coverage of ~7000 eukaryotic mRNA species (Table 2.1 and [309, 367]). For the first five primer combinations, DNA-free total RNA from each variant was reverse transcribed using the degenerate primer T12MG from the GenHunter RNA MAP kit (GenHunter Corp, Brooklyn, MA). Samples of each OCI LY8-C3P and C3-A11N reverse transcription (RT) were then amplified by low-stringency PCR (with a 40°C annealing step) in the presence of 35S- dATP (Dupont NEN, Boston, MA) and
employing the same T-specific primers used for RT in combination with various random 10-mers: AP1, AP2, AP3, AP4, or AP-5, also from GenHunter (Table 2.1). To show reproducibility of banding patterns, RNA from two independent subclones of each variant was used in independent RT reactions, and two separate PCR reactions were run for each representative RT reaction. A sample of each RT reaction, in which no reverse transcriptase was added, was subjected to the same PCR reactions and represented along with the rest of the samples (-RT) to control for contaminating chromosomal DNA. The resulting radioactive patterns of 3’ partial cDNA sequences were displayed on 6% denaturing polyacrylamide gels. Differential display with the rest of the primer combinations was similar except that the random primers themselves differed from the original RNA MAP primers: three one-base anchored T-specific primers, HT11A, HT11G, and HT11C, were used both in the RT and in combination with random 13-mers (H-AP1 to H-AP5) in the PCR (Table 2.1). These newer primer combinations have been reported to confer increased sensitivity and reduced redundancy [314, 361]. Sizes were determined by running known sequencing ladders adjacent to the displays, and cDNAs in the 150-500 bp range were selected based on reproducibility of banding across separate RT and PCR reactions. All displays were done using total RNA from variants in which differential expression of RAG1 and RAG2 was confirmed by Northern blotting prior to displays. cDNA bands differentially expressed are designated such that the first character represents the 10-mer primer, the second character the T-primer, and the third their order, from largest (first) to smallest (last).

For isolating candidate differentially-expressed cDNAs, bands of interest were excised from the polyacrylamide gels, eluted by boiling, and precipitated with the aid of glycogen as a carrier [309]. Reamplifications were done using the same primer set and PCR conditions except that the dNTP concentrations were 10 fold higher and no α-S35 dATP was added. cDNAs that failed to be reamplified in the first round of PCR were diluted and used as a template in a subsequent round of amplification. After PCR amplification, 30 μl of each of the final reaction products was electrophoresed on high percentage (2%) EtBr-stained agarose gels run
Table 2.1

Oligonucleotide primers used for DD RT-PCR$^1$.

<table>
<thead>
<tr>
<th>Oligo (dT) primers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>T12MG:</td>
<td>5'-TTTTTTTTTTTTTTMG-3', where M=equal mixture of A, C, and G</td>
</tr>
<tr>
<td>HT11A:</td>
<td>5'-AAGCTTTTTTTTTTA-3'</td>
</tr>
<tr>
<td>HT11C:</td>
<td>5'-AAGCTTTTTTTTTTC-3'</td>
</tr>
<tr>
<td>HT11G:</td>
<td>5'-AAGCTTTTTTTTTTG-3'</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arbitrary primers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Arbitrary 10-mers</td>
<td></td>
</tr>
<tr>
<td>AP1:</td>
<td>5'-AGCCAGCGAA-3'</td>
</tr>
<tr>
<td>AP2:</td>
<td>5'-GACCGCTTGT-3'</td>
</tr>
<tr>
<td>AP3:</td>
<td>5'-AGGTCGACCGT-3'</td>
</tr>
<tr>
<td>AP4:</td>
<td>5'-GGTCTCCAC-3'</td>
</tr>
<tr>
<td>AP5:</td>
<td>5'-GTTCGATCC-3'</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Arbitrary 13-mers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>H-AP1:</td>
<td>5'-AAGCTTGATTCGC-3'</td>
</tr>
<tr>
<td>H-AP2:</td>
<td>5'-AAGCTTGACTGT-3'</td>
</tr>
<tr>
<td>H-AP3:</td>
<td>5'-AAGCTTTGGTCAG-3'</td>
</tr>
<tr>
<td>H-AP4:</td>
<td>5'-AAGCGTAAACGTC-3'</td>
</tr>
<tr>
<td>H-AP5:</td>
<td>5'-AAGCAGGATTGCT-3'</td>
</tr>
</tbody>
</table>

$^1$ Differential display primers, as depicted in this table, include 4 oligo-dT$_M$N primers ("M" represents the degenerate position) and 10 arbitrary primers (AP1-AP-5 and H-AP1-H-AP5, indicating arbitrary 10-mers and arbitrary 13-mers, respectively). This results in a total of 40 primer combinations used for DD RT-PCR in these studies.
in 1X Tris/acetate/EDTA with appropriate low-molecular weight markers in order to resolve the amplicons. cDNAs for subsequent radiolabeling reactions were purified from agarose gels using the Qiaex kit (Qiagen).

Northern blot analysis

Aliquots of 10-20 μg of total RNA were electrophoresed under denaturing conditions and transferred onto a ζ-probe nylon membrane (Bio-Rad, Hercules, CA) by overnight capillary transfer. Membranes were cross-linked in a Stratalinker 2400 UV crosslinker (Stratagene, LaJolla, CA), pre-hybridized, hybridized to the appropriate probe, washed, wrapped in plastic wrap and exposed to BioMax MS X-ray film (Eastman Kodak, Rochester, NY) at -70°C. All technical procedures were according to the protocol supplied by the manufacturer of ζ-probe, excepting that multiple washing regimens, from non-stringent washes (two, 30 minute ζ-probe 5% SDS washes at 65°C) to highly stringent washes (multiple, ζ-probe 0.5% SDS washes at 70°C) were used so as not to miss any low-copy number transcripts. As with display gels, reconfirmations were performed at least twice using total RNA from independent subclones of each variant to exclude irreproducible differentially expressed bands. Where applicable, the intensity of the hybridization signals were quantitated by the ImageQuant phosphorimager software (Molecular Dynamics, Sunnyvale, CA).

Probes used in Northern blot analysis were generated by radiolabeling cDNAs with 32P dCTP (Dupont) using the Quickprime random hexamer labeling kit (Stratagene). The following cDNAs were used as probes: 1) a 0.9 kb human RAG1 coding region fragment generated by XhoI and HindIII digestion of a 6.6 kb cDNA fragment supplied by Dr. D. Schatz (Yale University School of Medicine, New Haven, CT); 2) a 1.0 kb cDNA probe specific for the human β-actin gene used as a loading control, obtained from Dr. N. Lassam (University of Toronto, Toronto, ON). Because DD probes typically comprised less than 500 bp of 3' UTR, sensitivity was increased by modifying the random hexamer labeling reaction such that the T-specific primer was included in the
labeling mix. Additionally, each probe was purified using Nic spin columns in order to minimize background and to verify that incorporated nucleotides for each radiolabeled DD amplicon was \( \geq 2 \times 10^7 \) counts per minute (cpm).

**Isolation of radiolabeled amplicons detecting single, differentially expressed transcripts from Northern blots**

To isolate amplicons specifically detecting differentially expressed transcripts from those detecting non-specific bands in Northern blots (representing the majority of situations in our screenings), the region of the membrane in which the differentially expressed transcript was located was determined using an RNA ladder size standard (Gibco BRL), and subsequently excised. The radiolabeled amplicon probe was eluted by boiling, and precipitated using glycogen as a carrier. The eluted amplicon was then reamplified following the same procedure as in the initial round of reamplification. The resulting PCR amplicon, if corresponding to the expected size as determined from the previous round of reamplification, was verified to yield specific, differential expression by using it as a radiolabeled probe through one further round of Northern blot reconfirmations. Subsequently, these were cloned into the PCRII vector using the TA Cloning Kit (Invitrogen Corp., San Diego, CA).

**Reverse transcriptase PCR for assessing RAG expression**

Total RNA samples (2 \( \mu \)g) from OCI LY8-C3P and C3-A11N were treated with DNaseI (Pharmacia) to remove contaminating genomic DNA and then reverse transcribed using Superscript II reverse transcriptase (Life technologies), 1 \( \mu \)g of random hexamer (Gibco BRL), 1 mM dNTPs and the supplied buffer. After RNaseH digestion (Pharmacia), 1/10 of synthesized first strand cDNA was used to amplify either RAG1, RAG2, or \( \alpha \)-tubulin transcripts by PCR using a Perkin Elmer-Cetus thermal cycler in a 100 \( \mu \)l final reaction volume. The PCR primers used included: RAG1A, 5' CAG CGT TTG GCT GAG CTC CT 3' (RAG1 sense primer); RAG1B, 5' GGC TTT CCA GAG AGT CCT CA 3' (RAG1 antisense primer); hR2A, 5' TTC TTG GCA TAC CAG...
CAG 3' (RAG2 sense primer located at nucleotides 32-49 of human RAG2 cDNA); hR2C, 5' CTA TTT GCT TCT GCA CTG 3'(RAG2 antisense primer located at nucleotides 207-224 of human RAG2 cDNA); Tub-5', 5' CAG GCT CAA TGT GGC AAC CAG ATC GGT 3'; Tub-3', 5' GGC GCC CTC TGT GTA GTG GCC TTT GCC CCA 3'. PCR conditions were as follows: 29 cycles of 45 seconds at 94°C, 45 seconds at 52°C, and 45 seconds at 72°C. The initial cycle included 5 minutes at 94°C denaturation, and the final cycle included 10 minutes at 72°C extension. After PCR amplification, 15 μl of each of the final reaction products were electrophoresed on a 1.5% agarose gel run in 1X Tris/acetate/EDTA buffer. RT-PCR products were then transferred onto ζ-probe nylon membrane (BioRad) by overnight capillary transfer. Membranes were then cross-linked using a Stratalinker 2400 UV crosslinker (Stratagene), and pre-hybridized. Hybridization was performed with the following specific [γ-32P] ATP end-labeled oligonucleotides: RAG1T, 5' AAG TAT AGG TAT GAG GGA A 3'; Tub-P, 5' ACC TGA GCG AAC AGA GTC CAT C 3'. After washing, membranes were exposed to BioMax MS X-ray film (Kodak). Experimental controls included amplification of RNA, to control for genomic DNA amplification, and amplification of α-tubulin transcripts to demonstrate presence of cDNA. Standard curves for quantitation purposes were generated through amplification of either the human RAG1 cDNA cloned in the pBluescript SK+ plasmid or the RAG2 cDNA cloned into the pBluescript SK+ plasmid. Results were validated by performing two independent experiments.

**DNA sequencing and sequence data analysis**

The primers for sequencing were synthesized using a PerSeptive 8909 automated DNA oligosynthesizer (PerSeptive Biosystems Inc., Framingham, MA) and purified by oligonucleotide purification cartridge chromatography (Applied Biosystems, Weterstadt, Germany). DNA sequencing of the differential display products was performed by the ACGT corporation on either separate strands and/or from the same strand by cycle sequencing, using the Thermo Cycle Sequencing Kit (Amersham, Arlington Heights, IL) on a LICOR 4000L automated sequencer (LICOR Corp.). Sequences were analyzed and verified using the Sequencher 3.0.1 DNA analysis
program (Gene Codes Corp., Ann Arbor, MI). Nucleotide sequences were analyzed for homology or identity with known sequences in nucleotide databases (EMBL, GenBank, DDBJ, dbEST) using the database similarity search algorithms FASTA, BLAST-N and BLAST-P from the GCG software package version 8 (Madison, WI). DNA sequences of cDNA clones isolated in our screens have been submitted to the GenBank database (accession nos. AFO80573-AFO80578).
2.4 RESULTS

2.4.1 Differential constitutive and inducible RAG expression in OCI LY8 variants

Our laboratory has previously extensively characterized differential constitutive RAG expression which accompanies spontaneous in vitro secondary rearrangements in the human mature B cell line OCI LY8-C3P [223]. Specifically, in this system, the parental cell line OCI LY8-C3P undergoes Igλ gene rearrangements on its productively rearranged Igλ allele, thereby giving rise to several independent variants. slg- variants, resulting from non-productive secondary Igλ rearrangements, were isolated from the OCI LY8-C3P cell line by fluorescence activated cell sorting (FACS) followed by limiting dilution analysis (Figure 2.1). Spontaneously arising sIg+Id- clones resulting from a further secondary rearrangement event were in turn isolated from the expanded sIg- populations by plating at limiting dilutions. Accompanying these consecutive secondary rearrangement events is increased constitutive RAG1 and RAG2 mRNA expression in both the slg- and sIg+Id- clones. Differential RAG expression in the OCI LY8 system has also been observed in a distinct situation. Increased levels of RAG1 and RAG2 mRNA result when slg+ variants are cross-linked with anti-μ or when slg+ and slg variants are treated with PMA+ionomycin [75]. For both the constitutive and inducible differential RAG1 expression in this cell line, I previously determined that steady-state RAG1 mRNA differences are mediated both transcriptionally and post-transcriptionally. Furthermore, both inducible and constitutive increases in RAG1 mRNAs appear to require de novo protein synthesis [75, 264].

Of these two systems, the constitutive differences between low and high-RAG expressing variants was chosen as the preferred system for identifying potentially novel genes co-expressing with the RAGs because differences in RAG1 and RAG2 total RNA were considerably greater and more stable than those observed with slg cross-linking mediated increases. Also, a large number of signaling-associated genes may be induced which are not necessarily associated with RAG when
Figure 2.1. Schematic summary of clonal relationship in OCI LY8 cell culture system between RAG$^{lo}$ parental clone OCI LY8-C3P and RAG$^{hi}$ variants C3-A11N and A8-6P. Arrows point to clonal variants generated through secondary rearrangements at the human $\lambda$ light chain in culture (For details, see Stiemholm and Berinstein, 1993; reference [223]). The C3-A8N variant is included because it gives rise to the A8-6P variant. C3-A8N was originally thought to be slg$^-$, but has been shown by our lab to express a receptor termed slg$\Delta$CL at its surface, comprised of a truncated $\lambda$ chain in association with $\mu$ chain at low levels (Stiemholm et al, 1995; reference [411]). Unlike C3-A8N, the C3-A11N variant cannot undergo further rearrangements.
comparing stimulated versus unstimulated conditions, resulting in more irrelevant candidates to screen out.

2.4.2 Quantitation of differential RAG expression in OCI LY8 clonal cell variants and rationale for using OCI LY8-C3P and C3-A11N in DD RT-PCR

I quantitated the constitutive differences in RAG expression in OCI LY8-C3P and its clonally-related variants, C3-A11N and A8-6P both by Northern analysis and by reverse transcriptase PCR. By Northern blotting, the increase in RAG1 mRNA signals in C3-A11N and A8-6P when compared to the parental clone OCI LY8-C3P was 20.8 fold and 16.4 fold, respectively, as quantitated by Phosphorimager analysis (Figure 2.2A). A similar increase in RAG1 and RAG2 transcripts in the C3-A11N variant by comparison with the parental clone was observed by semi-quantitative RT-PCR analysis (Figure 2.2B). Based on plasmid standard dilutions and total RNA yield calculations, I estimate that in this assay, OCI LY8 C3P contains an average of ~100 RAG1 and ~5 RAG2 mRNA copies/cell, whereas the C3-A11N variant contains an average of 1-2 x 10^3 RAG1 and ~50 RAG2 mRNA copies/cell. These estimated levels of RAG transcripts in C3-A11N are comparable to other RAG-enriched tissues or pre-B cell lines [62].

The above variants are well suited for approaches aimed at isolating differentially-expressed genes for two reasons. First, the low and high RAG-expressing OCI LY8 variants are clonally related. I reasoned that using such variants would increase the efficiency of isolating RAG-associated genes by decreasing the number of differentially-expressed genes associated with cell line differences. In this context, I have previously demonstrated that OCI LY8 variants do not exhibit alterations in general phenotypic or differentiation markers (such as in the cell surface markers CD10, CD19, CD20, CD38, B7-1, MHC I and MHC II) other than their RAG expression [223, 264]. Furthermore, there is no measurable difference in activation of early or late general signaling parameters amongst these variants (such as Ca^{2+} flux, anti-phosphotyrosine profiles, proliferation, c-fos mRNA induction, and alterations in CD25 and CD71 expression) [75, 264].
Figure 2.2. Quantitative comparison of RAG mRNA expression level between OCI LY8-C3P parental cell line and clonally related variants.

A. Northern hybridization of 10 μg total RNA employing a 0.9 kb human RAG1 cDNA fragment and a 1.0 kb human β-actin cDNA fragment as probes. OCI LY8-C3P is the low RAG-expressing parental diffuse large cell lymphoma cell line and C3-A11N and A8-6P are two clonally-related high RAG-expressing variants. The T lymphoblastoid cell line CEM is a negative control for RAG expression. B. RT-PCR analysis of RAG1 and RAG2 expression in OCI LY8-C3P and C3-A11N. The amplification product sizes for RAG1, RAG2, and α-tubulin are indicated. The controls include: no reverse transcriptase added in the cDNA synthesis step (-RT), α-tubulin as the internal standard, and dilutions of human RAG1 and RAG2 plasmid cDNA fragments as positive controls.
CEM
OC1
LYS-C3P
C3-A11N
AS-6P
107 copies
106 copies
105 copies
104 copies
103 copies
102 copies
101 copies
A second important criteria for using OCI LY8-C3P and C3-A11N to isolate differentially expressed genes is the stability of RAG gene expression in these variants. The relative differences in RAG expression, both between low and high RAG-expressing variants as well as between variant subclones is stable over time [223]. In contrast, subclones of other RAG-expressing cell lines have been reported to express variable levels of these gene transcripts over time [56]. The C3-A11N clone is particularly well suited because it has exhaustively rearranged both its alleles and therefore cannot undergo further IgA rearrangements in culture [223].

2.4.3 Selection approach for choosing candidate RAG co-expressing genes

As a primary screen for isolating genes co-expressed with RAG1, I used differential display PCR (DD RT-PCR) to comparatively fingerprint mRNA transcripts between OCI LY8-C3P and its non-rearranging variant C3-A11N. DD RT-PCR was performed on OCI LY8-C3P and C3-A11N using standard degenerate primer combinations [311, 361]. To reduce artifacts, for each set of primers, I used several internal controls including the use of independent subclones, reactions without reverse transcriptase, and multiple PCR reactions. To control for the possibility of residual genomic DNA contamination after DNaseI treatment, I also ran -RT controls adjacent to the rest of the samples. Furthermore, where “bunching” was observed in the upper part of the sequencing gel (which has been shown to potentially increase heterogeneity of DD products [314]) I also performed short and long electrophoresis. A representative differential display PCR gel is shown in Figure 2.3A and representative reamplification of differentially expressed bands are shown in Figure 2.3B (see Materials and Methods and Figure 2.3 legend). As described in Table 2.2, 22 reproducible differences between OCI LY8-C3P and C3-A11N were found in total (8 expressed in OCI LY8-C3P and 14 expressed in C3-A11N). Of these putative differentially expressed products, 14 differences were found to be expressed exclusively in one variant and not the other (Table 2.2).
Figure 2.3. Differential display analysis to detect differentially expressed bands in the low RAG-expressing clone OCI LY8-C3P and the high RAG-expressing variant C3-A11N.

A. Representative denaturing polyacrylamide gel on which $^{35}$S dATP-labeled PCR products from OCI LY8-C3P (designated-) and C3-A11N (designated+) are run adjacent to each other. Differential display gels using two primer combinations (HT11G+AP1, HT11C+AP5) are shown as examples. To minimize false positive bands as a result of contaminating chromosomal DNA, a control for each reverse transcriptase reaction, in which no RT was added was subjected to the same PCR reactions and represented along with the rest of the samples (-RT). For each cell line, duplicate PCR products from two independent RNA aliquots were analyzed. Only differences (denoted by arrows) which could be reproduced in two subclones of each variant and in duplicate PCR reactions were selected for further analysis. A sequencing ladder was also run in order to size cDNAs. Displays were performed using total RNA extracted from variants kept in culture <3 weeks after thawing of original frozen stocks, and differential expression of RAG1 and RAG2 was confirmed prior to performing displays by Northern blotting using a 0.9 kb coding region human RAG1 cDNA as an $\alpha$-P$^{32}$ dCTP-radiolabeled probe in Northern analysis. B. Reamplification of differentially expressed uncloned cDNA products i.e. designated amplicates, excised from gels using same primer sets and similar conditions used in original display PCR. The series using the T-primer HT11A is shown as an example. Reamplified amplicates were run on 2% agarose gels along with 1 $\mu$g of a φX174-HaeIII digest in order to estimate product amount and sizes. In cases where more than one amplicate was reamplified, the correct differentially expressed band confirmed by subsequent Northern blotting is indicated by an arrow.
Table 2.2
Summary of DD RT-PCR (1st screening) results for selecting genes co-expressed with human RAG1.

<table>
<thead>
<tr>
<th>Band</th>
<th>T-primer</th>
<th>Arbitrary Primer</th>
<th>Size of DD amplicate (bp)</th>
<th>Expression Pattern</th>
<th>Relative Expression ¹</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OCI LYS-C3P</td>
<td>OCI LYS-C3P C3-A11N</td>
</tr>
<tr>
<td>1A1</td>
<td>HT₁₁A</td>
<td>AP-1</td>
<td>310²</td>
<td>OCI LYS-C3P</td>
<td>+++</td>
</tr>
<tr>
<td>2A1</td>
<td>HT₁₁A</td>
<td>AP-2</td>
<td>485</td>
<td>OCI LYS-C3P</td>
<td>+++</td>
</tr>
<tr>
<td>2A2</td>
<td>HT₁₁A</td>
<td>AP-2</td>
<td>470</td>
<td>C3-A11N</td>
<td>++</td>
</tr>
<tr>
<td>2A3</td>
<td>HT₁₁A</td>
<td>AP-2</td>
<td>415</td>
<td>C3-A11N</td>
<td>++</td>
</tr>
<tr>
<td>2A4</td>
<td>HT₁₁A</td>
<td>AP-2</td>
<td>411</td>
<td>C3-A11N</td>
<td>+++</td>
</tr>
<tr>
<td>4A1</td>
<td>HT₁₁A</td>
<td>AP-4</td>
<td>321</td>
<td>OCI LYS-C3P</td>
<td>+++</td>
</tr>
<tr>
<td>4A2</td>
<td>HT₁₁A</td>
<td>AP-4</td>
<td>220</td>
<td>OCI LYS-C3P</td>
<td>+++</td>
</tr>
<tr>
<td>4A3</td>
<td>HT₁₁A</td>
<td>AP-4</td>
<td>150</td>
<td>OCI LYS-C3P</td>
<td>+++</td>
</tr>
<tr>
<td>5A1</td>
<td>HT₁₁A</td>
<td>H-AP-5</td>
<td>320</td>
<td>C3-A11N</td>
<td>-</td>
</tr>
<tr>
<td>5A2</td>
<td>HT₁₁A</td>
<td>H-AP-1</td>
<td>315</td>
<td>C3-A11N</td>
<td>-</td>
</tr>
<tr>
<td>1C1</td>
<td>HT₁₁C</td>
<td>H-AP-1</td>
<td>440</td>
<td>OCI LYS-C3P</td>
<td>+++</td>
</tr>
<tr>
<td>2C1</td>
<td>HT₁₁C</td>
<td>AP-2</td>
<td>290</td>
<td>C3-A11N</td>
<td>-</td>
</tr>
<tr>
<td>3C1</td>
<td>HT₁₁C</td>
<td>AP-3</td>
<td>420</td>
<td>C3-A11N</td>
<td>++</td>
</tr>
<tr>
<td>4C1</td>
<td>HT₁₁C</td>
<td>H-AP-4</td>
<td>430</td>
<td>OCI LYS-C3P</td>
<td>+++</td>
</tr>
<tr>
<td>5C1</td>
<td>HT₁₁C</td>
<td>AP-5</td>
<td>440</td>
<td>C3-A11N</td>
<td>-</td>
</tr>
<tr>
<td>5C2</td>
<td>HT₁₁C</td>
<td>AP-5</td>
<td>281</td>
<td>C3-A11N</td>
<td>+</td>
</tr>
<tr>
<td>5C3</td>
<td>HT₁₁C</td>
<td>AP-5</td>
<td>279</td>
<td>C3-A11N</td>
<td>++</td>
</tr>
<tr>
<td>1G1</td>
<td>HT₁₁C</td>
<td>AP-1</td>
<td>240</td>
<td>C3-A11N</td>
<td>++</td>
</tr>
<tr>
<td>2G1</td>
<td>HT₁₁G</td>
<td>H-AP-2</td>
<td>460</td>
<td>C3-A11N</td>
<td>++</td>
</tr>
<tr>
<td>2G2</td>
<td>HT₁₁G</td>
<td>AP-2</td>
<td>472</td>
<td>C3-A11N</td>
<td>++</td>
</tr>
<tr>
<td>3G1</td>
<td>T₁₅MG</td>
<td>AP-3</td>
<td>157</td>
<td>C3-A11N</td>
<td>++</td>
</tr>
<tr>
<td>3G2</td>
<td>HT₁₁G</td>
<td>AP-3</td>
<td>350</td>
<td>OCI LYS-C3P</td>
<td>+++</td>
</tr>
</tbody>
</table>

¹ relative levels of cDNA differential display bands in OCI LYS-C3P and C3-A11N. Symbols indicate the following: (+++)= maximal expression, (++)= estimated 3-10 fold reduction in expression, (+)= estimated >10 fold reduction in expression, (-)= no expression.

² approximate sizes of cDNA differential display bands based on DNA ladders run adjacent to display products on sequencing gel. Sizes for 2A4, 4A1, 5C2, 5C3, 2G2, and 3G1 cDNAs are exact based on sequencing data.
Because up to 85% of differential display amplicons can be artifactual [314]. I verified the expression pattern originally seen in DD RT-PCR by Northern analysis to exclude "false positives" (i.e. bands detecting transcripts of equivalent levels in OCI LY8-C3P and C3-A11N). This "secondary screening" was done by radiolabeling the purified, reamplified uncloned products isolated from the differential displays i.e. hereafter designated as amplicates, to probe total RNA from the original low and high RAG-expressing OCI LY8 variants. Nine of the 22 putative differentially expressed amplicates were reconfirmed by Northern blot analyses. (Figure 2.4 and Table 2.3). These nine reconfirmed amplicates detected 11 differentially expressed transcripts in all. The differentially expressed transcripts varied in size (from 400 bp to 10.6 kb) thereby pointing to the likelihood that independent differentially expressed transcripts had been isolated.

The exception to this was the 5C2/5C3 pair which hybridized to the same size differentially expressed transcripts as well as being found within several bp of each other on the sequencing gel used for display. The expected differences in RAG1 between variants was validated by sequentially hybridizing the same membrane strips with the radiolabeled RAG1 and β-actin cDNAs (Figure 2.4). Use of amplicates as probes in all cases yielded multiple non-specific bands, almost all of which were of more abundant message classes (see Table 2.3).

To extend my results from the secondary screens, a "tertiary screen" was performed. To identify amplicates hybridizing to transcripts with a RAG1 specific expression pattern, amplicates from the secondary screen detecting were used to probe total RNA from an array of human RAG+ and RAG- cell lines (Figure 2.5). RAG+ total RNA sources included another independently-derived high RAG-expressing OCI LY8 variant, A8-6P, and the pre-B cell lines 697 and Nal-1. RAG- total RNA sources included the human mature B cell lines Raji, Daudi, and OCI LY1, the mature T cell line Jurkat, and the non-lymphoid cell lines HeLa, U937, and K562. The expected differences in RAG1 between variants were validated by sequentially hybridizing the same membrane strips with the radiolabeled RAG1 and β-actin cDNAs. From this screen, I found that four transcripts exclusively expressed in C3-A11N were also found in other RAG-expressing cell lines assessed: a
Figure 2.4. Reconfirmation of differentially expressed amplicates between OCI LY8-C3P and C3-A11N by Northern blot.

A. Denaturing sequencing gels on which duplicate $^{35}$S dATP-labeled PCR products from two independent RNA aliquots from OCI LY8-C3P (designated -) and C3-A11N (designated +) are run adjacent to each other. Differentially expressed amplicates were first identified by DD RT-PCR as described in Materials and Methods. Only differences (denoted by arrows) which could be reproduced in two subclones of each variant and in duplicate PCR reactions were selected for further analysis. B. Representative Northern strips for differentially expressed amplicates 2A1, 2A3, 2A4, 4A1, 5C1, 5C2, 5C3, 2G2, and 3G1. Reproducible, differentially expressed amplicates were purified, radiolabeled, and used as probes to detect corresponding differentially expressed transcripts in OCI LY8-C3P and C3-A11N total RNA (denoted by arrows). The same blots were probed sequentially with RAG1 and β-Actin cDNAs. To minimize missing differentially expressed transcripts, amplicates were used at high specific concentrations ($\geq 5 \times 10^6$ cpm/ml), and multiple wash stringencies were employed. Representative blots have exposure times that vary considerably; all blots are shown under low-stringency wash conditions, as almost invariably, differentially expressed transcripts were not detected at higher stringency (see Materials and Methods). All differentially expressed transcripts detected under low stringency were present in at least two independent probings.
Table 2.3
Summary of Northern blot analysis (2° screening) for selecting differential display transcripts co-expressed with human RAG1.

<table>
<thead>
<tr>
<th>Probe</th>
<th>Estimated transcript size(s) detected (kb)</th>
<th>Expression pattern(s)</th>
<th>Relative difference in expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A1</td>
<td>4.0, 2.9, 1.8, 1.0</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>2A1</td>
<td>3.8, 1.7, 1.2</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>2A2</td>
<td>3.7, 2.2, 1.5</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>2A3</td>
<td>0.8</td>
<td>C3-A11N</td>
<td>qualitative</td>
</tr>
<tr>
<td>2A4</td>
<td>3.8, 2.0, 1.5</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>3A1</td>
<td>0.6</td>
<td>C3-A11N</td>
<td>qualitative</td>
</tr>
<tr>
<td>4A1</td>
<td>3.9, 1.0, 0.6</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>4A2</td>
<td>3.8, 2.5, 1.7, 1.0</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>4A3</td>
<td>3.5, 2.6, 1.6, 0.8</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>5A1</td>
<td>7.5, 3.7, 1.6, 1.1</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>5A2</td>
<td>7.1, 4.2, 2.1, 1.0, 0.8</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>1C1</td>
<td>7.5, 4.1, 2.2, 0.5</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>2C1</td>
<td>8.1, 3.6, 2.3, 1.0</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>3C1</td>
<td>7.1, 3.8, 2.4, 2.0, 1.1</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>4C1</td>
<td>7.2, 3.7, 3.2, 1.4</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>5C1</td>
<td>0.7</td>
<td>C3-A11N</td>
<td>4.5</td>
</tr>
<tr>
<td>5C2</td>
<td>1.0</td>
<td>C3-A11N</td>
<td>8.1</td>
</tr>
<tr>
<td>5C3</td>
<td>0.6</td>
<td>C3-A11N</td>
<td>4.3</td>
</tr>
<tr>
<td>1G1</td>
<td>3.8, 2.5, 1.7, 1.0, 0.9</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>2G1</td>
<td>3.6, 2.0, 1.6, 1.0, 0.5</td>
<td>non-specific</td>
<td>—</td>
</tr>
<tr>
<td>2G2</td>
<td>2.1</td>
<td>C3-A11N</td>
<td>12.9</td>
</tr>
<tr>
<td>3G1</td>
<td>5.0</td>
<td>C3-A11N</td>
<td>qualitative</td>
</tr>
<tr>
<td>3G2</td>
<td>3.8, 2.6, 1.8</td>
<td>non-specific</td>
<td>—</td>
</tr>
</tbody>
</table>

1 bands are denoted based on their presence in at least two independent probings and were calculated based on their relative positions to 28S and 18S rRNA.
2 total RNA from the low RAG-expressing line OCI LY8-C3P and the high RAG-expressing variant C3-A11N were assessed by northern analysis and probed with reamplified, uncloned differentially-expressed products i.e. amplicates.
3 relative differences were calculated by measuring intensities of the given probe and β-actin by phosphorimager analysis and calculating probe/β-actin ratios.
Figure 2.5. Tertiary screens showing expression of transcripts detected by differentially expressed amplicates in unrelated RAG- and RAG+ cell lines.

Expression pattern of differentially expressed transcripts in RAG+ and RAG- sources using the reconfirmed the differentially expressed amplicates 2A4, 4A1, 5C2/5C3, 2G2, and 3G1 as radiolabeled probes in Northern analysis. Arrows indicate the position of differentially expressed transcripts. The same blots were then sequentially hybridized with a 40-457 bp human RAG1 cDNA probe and a 1.2 kb EcoRI-generated β-actin cDNA probe. The exposure times were as follows: 7d for 2A4, 4d for 4A1, 10d for 5C2/5C3, 14d for 2G2 and 3G1, 16 h for RAG1, and 12 h for β-actin.
5.0 kb mRNA detected with the 3G1 amplicate, ~600 bp and ~1.0 kb mRNAs detected with the 5C2 and 5C3 amplicates, a ~10.6 kb mRNA detected with the 2A4 amplicate, and a ~2.1 kb mRNA detected with the 2G2 amplicate (Figure 2.5). One of the two transcripts differentially expressed in OCI LY8-C3P, a ~2.3 kb transcript detected with the 4A1 amplicate, was also found predominantly in other non-RAG expressing cell lines.

2.4.4 Purification of single amplicons from heterogeneous amplicate mixtures

Based on the multiple hybridizing bands seen in our secondary and tertiary screening results, it was evident that the 22 DD amplicates either corresponded to genes with alternative splicing patterns or conversely, were mixed and detected distinct genes (Figures 2.4 and 2.5; Table 2.3). I initially attempted to clone these amplicons directly into PCRII and to subsequently use these cloned products as probes. However, with this approach, usually only one of the non-specific, more abundant transcripts was obtained as determined by Northern blot analysis. Thus, in order to sequence the appropriate differential expressed transcript, it was necessary to isolate the unique amplicons detecting the differentially expressed transcripts from irrelevant contaminating amplicons. This was done by eluting radiolabeled cDNA from the region in the Northern blots in which the transcripts of interest were estimated to be located, and reamplifying this cDNA for further manipulation. Reamplified products were then electrophoresed on agarose gels and purified. To confirm that selection for the correct amplicon occurred, these products (designated amplicons) were cloned into the PCRII vector and used as radiolabeled probes in Northern analysis of total RNA from OCI LY8-C3P and C3-A11N. Purification of the correct amplicon should be confirmed by single, differentially expressed transcripts. As an example, Figure 2.6A shows purification of the unique amplicon corresponding to the differentially expressed transcript from the 5C2 DD amplicate, the latter a product that yielded two non-specific transcripts and two differentially expressed transcripts in our secondary and tertiary screens (Figures 2.4 and 2.5). In the case of the 5C2 and 5C3 amplicates, four cDNAs were isolated (two differentially expressed transcripts/amplicate). Both the 600 bp and 1.0 kb bands were excised and eluted from these gels
Figure 2.6. Representative isolation of unique amplicons from heterogeneous DD RT-PCR amplicate mixtures.

A. Original Northern blot from the secondary screening showing two differentially expressed transcripts and two non-specific transcripts detected by the 5C2 amplicate. B. Agarose gel analysis of the reamplified products (each 182 bp) derived from the eluted, radiolabeled 5C2 cDNAs specifically detecting 1.0 and 600 bp differentially-expressed transcripts (referred to as 5C2/5C3-1.0 and 5C2/5C3-0.6 amplicons, respectively). The unique amplicate products hybridizing to the differentially-expressed transcripts shown in A were cut out of the appropriate region in the Northern membrane, eluted and reamplified under the same conditions used for reamplifications from polyacrylamide gels in the original reconfirmation steps. C. Representative Northern blots of 10 μg total RNA from OCI LY8-C3P and C3-A11N that have been sequentially hybridized: first, with either the 5C2/5C3-1.0 or 5C2/5C3-0.6 reamplified and purified amplicons, and then with a human β-actin cDNA probe. The reamplified, purified amplicon detect only the differentially expressed transcript(s) from which they were purified. The 5C2/5C3-1.0 and 5C2/5C3-0.6 amplicons and all other amplicons isolated from amplicates after tertiary screens were subcloned into PCRII prior to sequencing.
probe: 5C2 amplicate

- 1.0 kb (contains 5C2/5C3-1.0 amplicon)
- 600 bp (contains 5C2/5C3-0.6 amplicon)
and reamplified with the same primer sets as originally used. The re-amplified amplicon purified from the 5C2 amplicate are shown as an example in Figure 2.6B. The 600 bp and 1.0 kb transcripts are from distinct genes (not alternative splicing) because each resulting unique, purified amplicon detected single 600 bp or 1.0 kb bands, respectively when used as a radiolabeled probe in Northern analysis (Figure 2.6C).

2.4.5 Sequence analysis of purified amplicons

As expected, the 5C2 and 5C3 amplicates each contained the identical two amplicons. This was fully consistent with the identical pattern of differentially expressed transcripts detected by the 5C2 and 5C3 amplicates in secondary and tertiary screens (Figures 2.4 and Table 2.2), and also consistent with the fact that they were found in almost identical regions of the initial DD RT-PCR gel (Table 2.1). I have therefore labeled the two unique amplicons detecting the ~600 bp and 1.0 kb transcripts 5C2/5C3-0.6 and 5C2/5C3-1.0, respectively. The sequences of the purified, unique amplicons are shown in Table 2.4. The sequences include matches to: a portion of the human RAG2 exon 2 coding region (corresponding to the purified 2G2 amplicon; accession no. AFO80577), the human BSAP (Pax-5) complete cDNA (corresponding to the purified 2A4 amplicon; accession no. AFO80573), and two human expressed sequence tags (amplicons 4A1 and 5C2/5C3-1.0; accession nos. AFO80574 and AFO80575, respectively). Two sequences (amplicons 5C2/5C3-0.6 and 3G1; accession nos. AFO80576 and AFO80578, respectively) did not have database sequence matches.

Based on these partial sequences, several important technical points can be made regarding the DD RT-PCR technique itself. Firstly, under the standard low stringency conditions used in our system, the primers often did not anneal to the theoretically intended target sequence i.e. the polyadenylated 3’ region of differentially-expressed transcripts (Table 2.4). For example, based on the sequence of the BSAP cDNA matching to the 2A4 amplicon, the arbitrary decamer AP-2 was used as both the sense and the antisense primer to amplify a portion of the BSAP coding
Table 2.4
Sequences of amplicons detecting human RAG1 co-expressing transcripts.

<table>
<thead>
<tr>
<th>Purified cDNA</th>
<th>accession no.</th>
<th>Corresponding genbank sequence</th>
<th>% homology to corresponding genbank sequence</th>
<th>Corresponding protein function</th>
<th>Corresponding transcript size/ expression pattern/ chromosomal location</th>
</tr>
</thead>
<tbody>
<tr>
<td>2A4</td>
<td>AFO80573</td>
<td>Human BSAP (Pax-5) mRNA, complete cDNAs (bp 258-668); accession no. M96944</td>
<td>99</td>
<td>Transcription factor involved in B cell development</td>
<td>10.6 kb transcript expressed in early B cells, developing neuronal cells, and testis</td>
</tr>
<tr>
<td>4A1</td>
<td>AFO80574</td>
<td>Human ESTs R40071, N54804 + 19 others</td>
<td>98</td>
<td>?</td>
<td>Corresponding EST isolated from brain, breast, colon, eye, foreskin, ovary, placenta, uterus cDNA libraries; maps to human chromosome 12 (D12S328)</td>
</tr>
<tr>
<td>5C2/SC3-1.0</td>
<td>AFO80575</td>
<td>Human ESTs H54661, AA280289, H40149, H40159, H94982, AA214529</td>
<td>99</td>
<td>?</td>
<td>Corresponding ESTs isolated from tonsil, fetal spleen cDNA libraries; map to human chromosome 11 (D11S1357-D11S1765)</td>
</tr>
<tr>
<td>5C2/SC3-0.6</td>
<td>AFO80576</td>
<td>No matches</td>
<td>—</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>2G2</td>
<td>AFO80577</td>
<td>Human RAG2 exon 2 + flanking genomic sequence (bp 2461-2872); accession no. M94633</td>
<td>99</td>
<td>Enzyme that activates V(D)J recombination in conjunction with RAG1</td>
<td>2.1 kb transcript expressed in early B and T lymphocytes; maps to human chromosome 11 (D11S907-D11S935)</td>
</tr>
<tr>
<td>3G1</td>
<td>AFO80578</td>
<td>hBRAG complete cDNA (bp 4225-4381); accession no. AFO26477</td>
<td>100</td>
<td>RAG1 regulation ? B-cell development?</td>
<td>5.0 kb transcript expressed at highest levels in early and mature B cells; low levels of expression in several other tissues; maps to human chromosome 10q26 (D10S244E)</td>
</tr>
</tbody>
</table>

129
region rather than the 3’ untranslated sequence. This finding is consistent with reports that many DD products actually amplify internal sequences rather than polyadenylated regions [368]. I also found that at least four mismatches were tolerated with the arbitrary decamer in all six of our sequences, and that these mismatches were predominantly clustered in the 5’ end of the decamer. The 5’ region of the decamer being preferential for mismatches is consistent with previous analysis [367], although the frequency with which mismatches were tolerated (both for different sequences and within the same sequence) was higher in our study. One possible reason for this could be differences across studies in the stringency of the DD RT-PCR parameters. Finally, in 3/5 purified amplicons (2G2, 4A1, and 5C2/5C3-0.6) the polyA-anchoring primer was also found to tolerate up to four mismatches and hence did not necessarily amplify the very 3’ polyA-ends of differentially expressed transcripts. In the case of RAG2 (2G2 amplicon) for example, an adenine-rich stretch in the RAG2 coding region resembling a polyadenylated tail was detected by the poly-A anchoring primer.
2.5 DISCUSSION

Lymphocyte differentiation is a complex series of events mediated by the expression of a number of lineage- and differentiation-restricted genes. I hypothesized that using DD RT-PCR to assess a unique pair of clonally related variant cell lines with differential expression of RAG transcripts could be used to isolate developmentally and B lineage-restricted genes by virtue of their co-expression with human RAG1 in B cell lines. There are several examples of pre-B and/or pre-lymphocyte-specific genes with a RAG-like expression pattern including the early B-cell specific transcription factors N-myc, LEF-1, and EBF, the V(D)J -modifying enzyme TdT, the cell surface marker CD10, and components of the pre-B cell receptor VpreB and λ5.

The six partial sequences of RAG1 co-expressing genes identified in this study demonstrate the power of my approach. One of the amplicons identified, 2G2, corresponds to the 3' end of human RAG2 exon 2 (Table 2.4), and the transcript size of ~ 2.1 kb detected in tertiary screens (Figure 2.3) is consistent with the known size of the predominant human RAG2 transcript [223]. RAG2 is one of the differentially-expressed genes expected to be picked up between these two variants, and thus serves as a good internal control. Because I did not use all possible primer combinations, it is likely that the more abundant RAG1 message in my screens was undetected because the primer sets employed did not cover this particular transcript.

Interestingly, another of the differentially expressed amplicon sequences (2A4; Table 2.4) has complete identity to an internal region of the human transcription factor BSAP (Pax-5) cDNA, a B-cell specific member of the paired domain Pax family of transcription factors required not only in B cell differentiation [369], but also for commitment to the B cell lineage [370]. The 10.6 kb differentially-expressed transcript detected by 2A4 in our secondary and tertiary screens in RAG+ sources (Figures 2.4 and 2.5) is consistent with the predicted size of the Pax-5 mRNA [371]. Although it has been previously established that BSAP is expressed in pro-B, pre-B, and mature B cells as well as the CNS and testis, I show here that human BSAP expression correlates with
human RAG1 expression, suggesting a potential functional link between these two genes. Supporting these observations indirectly are the phenotypes of BSAP and E2A-deficient mice. Mice with disruptions in the Pax-5 gene have an arrest in B cell development very similar to that of RAG knockout phenotype with the exception that residual heavy chain rearrangements can still be detected [369]. Furthermore, the E2A knockout mouse, which is also blocked at the pro-B cell stage of differentiation, has reductions in both the RAG1 and Pax-5 transcripts [207]. Because BSAP has an important role in B-cell differentiation, and as such has been demonstrated to interact with several B-cell specific gene promoters (for example CD19, λ5, and mb-1) [372, 373], it is therefore tempting to speculate that BSAP may be involved in inducing or enhancing the tissue-specific expression of RAG1 mRNA. Although the sequences of the core RAG1 and RAG2 promoter regions by themselves have been demonstrated not to be important for developmental and tissue-specific expression of RAG transcripts [262, 266], Lauring and Schlissel have recently found that the mouse RAG2 promoter can be transactivated in a tissue specific fashion through binding of a conserved BSAP site within the RAG2 core promoter [271]. Other elements outside of the promoters, perhaps including those containing BSAP motifs, may also be required. It will be interesting to test if stable transfection of the human BSAP cDNA in various human cell lines has a role in the B cell-specific regulation of RAG expression. This issue is addressed further in the general discussion in light of the recent studies by Lauring and Schlissel.

My screens also yielded two amplicons (4A1 and 5C2/5C3-1.0) with matches to human ESTs. Consistent with the differential expression of the transcript detected by 4A1 in the low RAG-expressing variant OCI LY8-C3P (in secondary screens) and with the inverse correlation with RAG expression in tertiary screens, the 4A1 amplicon matches several ESTs derived from non-lymphoid sources (Table 2.4). Conversely, the cDNA (5C2/5C3-1.0) detecting a 1.0 kb transcript with a similar expression pattern to RAG has a near-identical match to several ESTs from germinal centre B-cell and splenic-derived cDNA libraries (Table 2.4). Although this EST maps to chromosome 11, like the RAG locus, it is not physically proximal to the RAG genes.
The other two amplicons (5C2/5C3-0.6 and 3G1) were found to have no homology to any database sequences. As with 5C2/5C3-1.0, these two amplicons of unknown sequence detected transcripts that were found in the high RAG-expressing clone C3-A11N. I have cloned the corresponding full-length cDNA of one of these unknown sequences (3G1). This cDNA (accession number AF026477) encodes a novel gene I have called hBRAG (human B-cell RAG-Associated Gene), which encodes a type II transmembrane glycoprotein unrelated to other known type II protein-encoding multigene families. The molecular and biochemical characterization of this gene will be discussed in Chapters 3 and 4, respectively.

DD RT-PCR was a new methodology when I first began and in the course of this work I encountered two major technical limitations which had not been addressed in the literature. The first limitation is the high frequency at which false positive artifacts arise using standard DD RT-PCR methodology [314]. Such artifacts are not only problematic in that they may actually be masking bona fide differentially expressed transcripts in the same region of Northern blots, but they also make screening for differentially-expressed genes extremely inefficient. In my DD RT-PCR screens, 14 of 22 of DD amplicates were "false positives", despite the fact that clonally related cell variants were used along with all appropriate controls. While it has been reported that a large fraction of artifacts can be further reduced by increasing the stringency of the PCR annealing conditions or using more specific (less degenerate) primers [367], it has also been documented that such increased stringency parameters results in the out-competition of lower abundance class transcripts (to which most developmentally-regulated genes belong) by more-abundant, non-specific transcript classes [314]. During the course of optimizing the technique for our system, I found that while either increasing the stringency of the standard annealing step or reducing cycles numbers reduces contaminants that produce strong non-differential bands, it also compromises the detection of reproducible differentially expressed transcripts detectable by DD RT-PCR. For example, as mentioned, many of the amplicates that proved interesting had a large number of mismatches at the 5' end. These would undoubtedly have gone undetected under high-stringency
conditions. Instead, by increasing the screening stringency in steps subsequent to the DD RT-PCR itself i.e. performing two separate rounds of Northern blotting, I eliminated 17 of 22 amplicons found in the original DD RT-PCR, while at the same time testing the amplicates on other RAG* sources (Figures 2.4 and 2.5).

The second major limitation of the DD RT-PCR technique was that almost every one of the original amplicates was a mixture of amplicons, most of which detected non-specific genes. Attempts to resolve different species by running both long and short gels were not successful: I still found heterogeneous transcript species under low stringency Northern blotting conditions (Table 2.3 and Figure 2.4). As noted by others, my results indicate that prior to subcloning and sequencing, purification of single cDNAs (at least under the standard DD RT-PCR conditions used in our assays) was a necessity. I have shown that direct elution of the amplicon of interest from the Northern membrane is an effective way of doing this (Figure 2.6). Other reports have also discussed ways to purify single cDNAs from heterogeneous DD amplicons [374-380]. In all, my approach for eliminating artifactual cDNA products allowed us to eliminate ~75% of all DD amplicons and an additional ~ 75% of cDNAs within each amplicon of interest. A flowchart summarizing my overall scheme for selecting RAG co-expressing cDNAs is shown in Figure 2.7.

How efficient is the DD RT-PCR technique at detecting developmentally and lineage-restricted genes? One report supports the notion that DD RT-PCR is biased towards identifying a small number of intermediate and high message class transcripts (≥300 and 12000 mRNA copies/cell, respectively) [381]. Because 90-95% of eukaryotic mRNA species are estimated to belong to the rare message class (<50 copies/cell) [315], this would imply DD RT-PCR misses the majority of mRNAs. Conversely, another study argues for the exquisite sensitivity of DD RT-PCR in detecting rare message transcripts [368]. In my study, the fact that DD RT-PCR could identify RAG2, a low abundance transcript in C3-A11N (see Figure 2.2B), supports the latter report. Furthermore, the other differentially expressed transcripts detected in secondary and tertiary screens may also belong to the rare message class, as suggested by the relatively faint signal by
Figure 2.7. Summary of overall selection scheme for isolating candidate RAG co-expressing human genes. The fraction of cDNAs at each level of screening that were selected for the next stage of analysis are denoted by horizontal arrows.
comparison to non-specific, high-abundance transcripts within the same amplicon (Figures 2.4 and 2.5).

My experiments have confirmed that DD RT-PCR can be utilized to isolate genes with a very specific pattern of expression i.e. expressed either directly or indirectly with the RAG genes. As well as detecting RAG2 itself, I isolated BSAP (Pax-5), previously shown by others to be associated with lymphoid recombination, although not directly with RAG expression itself. Other novel cDNAs have also been cloned, and I have recently demonstrated that one of these (hBRAG) may itself regulate the expression of human RAG1. The sequences of other cDNA clones are currently being analysed and the relevance of the isolated genes to the RAGs and to early B cell development will be determined. I have shown that refinements to DD RT-PCR can yield a powerful approach to the isolation of specific and novel lineage- or developmentally-restricted genes.
CHAPTER 3

Molecular cloning and characterization of hBRAG, a Novel B-cell Lineage cDNA encoding a Type II Transmembrane Glycoprotein Potentially Involved in the Regulation of Recombination Activating Gene 1 (RAG1)
Acknowledgements

Much of the material in this chapter is published in the article entitled *hBRAG, a Novel B-cell Lineage cDNA encoding a Type II Transmembrane Glycoprotein Potentially Involved in the Regulation of Recombination Activating Gene 1 (RAG1)* in the European Journal of Immunology, volume 28(9), pages 2839-2853, by Laurent K. Verkoczy, Philip A. Marsden, Neil L. Berinstein (1998). I performed all experiments described in this chapter except for the chromosomal mapping of hBRAG. The FISH mapping was done by Barbara Beatty of the CGAT FISH Mapping Resource Facility and the radiation hybrid studies were done in the laboratory of Dr. Phil Marsden. The ACGT Corporation did a portion of the sequencing. Reinhart Reithmeier and Florence Tsui assisted me in the sequence analysis of the hBRAG cDNA and David Hogg assisted me with initial database searches. This work was supported by the National Cancer Institute of Canada (NCIC grant # 7286).
3.1 ABSTRACT

The Differential Display RT-PCR (DD RT-PCR) technique was used to identify novel cDNAs detecting mRNA transcripts co-expressed with human recombination activating gene-1 (RAG1). A 5.0 kb transcript detected by the Differential Display amplicon 3G1 was found to correlate strongly with RAG1 mRNA expression in various human cell lines. Subsequent screenings of a pre-B cDNA library with 3G1 led to the identification of a complete cDNA that we termed hBRAG (human B-cell RAG-Associated Gene). The hBRAG cDNA encodes a 503 amino acid (aa) protein with no known homology to any nucleotide or protein sequence. The predicted molecular mass of 55 kD was confirmed by in vitro translation. Based on sequence analysis, the predicted open reading frame encodes a type II transmembrane spanning glycoprotein with the N-terminal 81 aa in the cytoplasm, a 17 aa transmembrane domain, and a C-terminal 405 aa extracellular domain with four potential N-glycosylation sites. Northern blot analysis indicated a close association of the 5.0 kb hBRAG mRNA transcript with RAG1 in numerous human pro-B, pre-B and mature B cell lines assessed, but not in human T cell lines. In human tissues, hBRAG is expressed at highest levels in B cell-enriched tissues, but is not expressed in fetal or adult thymus. Southern blotting analysis revealed that this gene is single copy in the human genome, is highly conserved across vertebrate species, and is likely not a multigene family member. The hBRAG gene was localized to the long arm of chromosome 10 (10q26). Transfection of the full length hBRAG cDNA increased levels of human RAG1 transcripts in the B cell line OCI LY8-C3P, but not in the non-lymphoid line K562, suggesting a B cell-specific role for the hBRAG product in regulating RAG expression.
3.2 INTRODUCTION

Central to V(D)J recombination are the products of the recombination Activating Gene 1 (RAG1) and Recombination Activating Gene 2 (RAG2) locus [43, 44]. Several lines of evidence have demonstrated their pivotal role in this process. Firstly, they are necessary and sufficient to induce V(D)J recombination of exogenous recombination substrates in non-lymphoid cell lines and their expression pattern correlates precisely with recombinase activity [38]. Additionally, mice deficient in either gene product fail to generate mature B or T lymphocytes accompanied with the retention of Ig and TCR alleles in germline configuration [50, 83]. Finally, it has recently been demonstrated using an in vitro cell-free system that the RAG proteins play a direct role in the specific recognition and cleavage of V(D)J substrates [170, 362] as well as stimulation of coding joint formation [158].

The expression of RAG1 and RAG2 is precisely regulated. For the most part, RAG1 and RAG2 are expressed concordantly and in a stage-specific manner only in cells of the lymphocyte lineage [37]. Briefly, a first "wave" of RAG expression is detected in committed B and T progenitor lymphocytes at the point where Ig µ and TCR β and δ loci undergo V(D)J rearrangement [39]. The second "wave" corresponds to V(J) rearrangement of the TCR-α [41] and Igκ and λ loci [53]. Recently, third and fourth "waves" of RAG expression have been observed by several groups in Ig* B cells [62, 71, 72, 75, 223, 248] and TCR* thymocytes [222, 224].

At the molecular level, it is not well understood what trans-acting factors are responsible for controlling the temporal and tissue-specific pattern of expression of RAG1 and RAG2. Additionally, the growing list of putative recombinase components whose genes have been cloned [145, 382-386] combined with the emerging knowledge that V(D)J recombination is a multi-step process [387] indicates the participation of additional, as yet unidentified factors.

I sought to isolate and characterize novel genes that may fall into one of these two aforementioned categories by virtue of their co-expression with RAG1 mRNA. In this chapter, I report the cloning and initial characterization of a gene, hBRAG (human B-cell RAG-Associated Gene), that was
obtained by first isolating its 3’ end through the differential display approach described in Chapter 2, and subsequent cloning of the corresponding full-length cDNA. Based on the hBRAG expression pattern, proposed structure, and results of stably transfecting its full-length cDNA into OCI LY8-C3P, I suggest that the product encoded by the hBRAG gene may be involved in the B cell-specific regulation of the RAGs and therefore plays an important role in B cell development.
3.3 MATERIALS AND METHODS

Cell lines and tissue culture

The human B cell line OCI LY8-C3P (μλ+) is a single-cell clone from the OCI LY8 cell line (IgM+/IgD/CD10+/CD19+/CD20+/CD38+), originally established from a patient with a B lymphoid large cell lymphoma [365]. The OCI LY8 variants have been detailed elsewhere [223]. Other human mature B cell lines used included the diffuse large cell lymphoma lines OCI LY1 and OCI LY2, and the Burkitt's lymphoma lines Daudi and GM-607. Human pro-B cell lines included REH (RAG1+, μ+, CD19+) and PB-207 (RAG1; λ5+, μ+); human pre-B cell lines included Nal-1, PB-697 (RAG1; λ5+, μ+) and SMS-SB (pre-B acute lymphocytic leukemia). The latter two were kindly provided by Max Cooper (University of Alabama, Birmingham, AL). T-lineage lines used in our studies included the pre-T lines Molt-4 and CEM-6, (gifts from Max Cooper) and the mature T line Jurkat. Non-lymphoid cell lines used in our studies include K562 (human erythroleukemia, ATCC CCL 243) and HeLa (ATCC, human epithelial carcinoma). The U937 (pro-monocytic) and HUVEC (Human Umbilical Vein Endothelial Cell) cell lines were kindly provided by Michelle Letarte (Hospital for Sick Children, Toronto, ON), the PC31 (prostate-derived) cell line was kindly provided by Ed Conway (University of Toronto, Toronto, ON) and a Mcl17 (melanoma-derived) cell line was kindly provided by Norm Lassam (University of Toronto). All cells were routinely cultivated at 37°C/5% CO₂ in RPMI 1640 medium (ICN, St-Laurent, QUE.), supplemented with 10% Fetal Calf Serum (FCS; Hyclone, Logan, UT), 2mM L-glutamine, 100 U/ml penicillin, and 100 μg/ml streptomycin (Gibco BRL, Grand Island, NY).

RNA sources and isolation

Total RNA was extracted from tissues or cells in log-phase growth (5x10⁵ cells/ml) by the single-step guanidium thiocyanate phenol-chloroform extraction procedure [366]. For normal human bone marrow, tonsil and peripheral blood leukocytes (PBL), total RNA was extracted using the TriZol kit according to the supplier's instructions (Gibco BRL). Total normal human thymus RNA
was kindly provided by David Hogg (University of Toronto). Poly(A)$^+$ enrichment of total RNA was performed using the Oligotex mRNA kit according to the supplier's instructions (Qiagen).

**Differential display and isolation of candidate differentially expressed cDNAs**

Differential display PCR (DD RT-PCR) was performed on DNA-free total RNA from OCI LY8-C3P and C3-A11N as previously described in section 1.3.4.1 of Chapter 1 (General Introduction) and in the Chapter 2 Materials and Methods.

**Northern blot Analysis**

10-20 μg total RNA or 2 μg poly(A)$^+$ RNA from human cell lines and tissues was electrophoresed under denaturing conditions and transferred onto a $\zeta$-probe nylon membrane (Bio-Rad, Hercules, CA) by overnight capillary transfer. Membranes were then cross-linked in a Stratalinker 2400 UV crosslinker (Stratagene, LaJolla, CA), pre-hybridized, hybridized to the appropriate probe, washed, wrapped in plastic wrap and exposed to BioMax MS X-ray film (Eastman Kodak, Rochester, NY) at -70°C. All technical procedures were according to the manufacturer’s protocol. Where applicable, hybridization signal intensity was quantitated with ImageQuant phosphorimager software (Molecular Dynamics, Sunnyvale, CA). The expression of the hBRAG transcript in tissues was also studied using a MasterBlot Multiple Tissue Northern (Clontech, Palo Alto, CA) which contained 500 ng of poly(A)$^+$ RNA/square extracted from 50 human tissues and organ types from normal, pooled sources. Equal loading of these dot blots was verified by the manufacturer with eight housekeeping gene controls. The filter was hybridized overnight at 65°C in Express Hyb Hybridization solution (Clontech) using appropriate $^{32}$P-labeled cDNAs as probes. The filter was washed and stripped between probings according to the manufacturer’s instructions except that an additional wash was carried out in 0.1X SSC and 0.5% SDS, at 55°C for 20 minutes.

Probes were generated by radiolabeling cDNAs with $^{32}$P dCTP (Dupont) using the Quickprime random-hexamer labeling kit (Stratagene) and were purified using Nic spin columns (Pharmacia).
The following cDNAs were used as probes: a 0.9 kb human RAG1 coding region fragment generated by Xhol and HindIII digestion of a 6.6 kb cDNA fragment supplied by David Schatz (Yale University School of Medicine, New Haven, CT); a 1.0 kb cDNA probe specific for the human β-actin gene, obtained from Norm Lassam; a 1.9 kb cDNA probe (8-3, 5') generated by EcoRI digestion of Nalm-6 cDNA clone 8-3 spanning the entire coding region of the human hBRAG gene; and a 1.5 kb cDNA probe (25108-5') generated by NotI and HindIII double digestion of the 5' end of human Expressed Sequence Tag (EST) clone 25108 (Research Genetics, Alabama, USA) specific for the 3' hBRAG untranslated (UTR) region.

cDNA library construction and screenings

The C3-A11N cDNA library was prepared as follows: twice selected C3-A11N poly(A)* RNA was isolated by affinity chromatography on oligo-d(T) cellulose columns (New England Biolabs, Beverly, MA). Double stranded size-selected cDNA which had EcoRI and Xhol restriction sites on the ends was synthesized from 10 µg of C3-A11N poly(A)* RNA using the ZAP Express cDNA synthesis kit (Stratagene). The cDNAs were ligated unidirectionally into the EcoRI/Xhol λ ZAP Express vector cloning sites, and packaged using the GigaPack® GoldII extract (Stratagene). For initial screening, 1x10^6 plaque forming units (PFU) (20,000 PFU/150 mM plate) were plated, duplicate colony lifts were prepared on Optitran supported nitrocellulose filters (Schleicher and Schuell, Keene, NH), and the λ phage was fixed by UV irradiation crosslinking using a Stratalinker 2400 UV crosslinker (Stratagene). Filters were prehybridized and hybridized in 6X SSC containing 5X Denhardt’s, 10 mM NaH₂PO₄ (pH 6.5), 1 mM EDTA, 0.75% SDS, and 100 µg/ml salmon sperm DNA. Prehybridization was carried out at 65°C for 3h. Hybridizations were carried out overnight at 65°C using the cloned 3G1 differential display amplicon radiolabeled with the Quickprime random priming kit (Pharmacia) as a probe (6x10⁶/ml or 1.2x10⁹ cpm total). Filters were washed twice in 1X SSC containing 10 mM NaH₂PO₄ (pH 6.5), 1 mM EDTA and 0.5% SDS for 30 minutes, twice in 0.5X SSC containing 10 mM NaH₂PO₄ (pH 6.5), 1 mM EDTA and 0.5% SDS for 30 minutes, twice in 0.2X SSC containing 5X Denhardt’s, 10 mM
NaH₂PO₄ (pH 6.5), 1 mM EDTA and 0.5% SDS for 30 minutes, and once in 0.1X SSC containing 5X Denhardt's, 10 mM NaH₂PO₄ (pH 6.5), 1 mM EDTA and 0.5% SDS for 15 minutes. Positive clones were selected, titered, and subjected to secondary and tertiary rounds of screening under identical conditions. The cDNA inserts in positive clones were excised from the Zap express phage vector arms *in vivo* by ExAssist helper phage-mediated circularization into the pBK phagemid (Stratagene). cDNA inserts were titered to single phagemid colonies for miniprep analysis of insert sizes. The 3' ends of cDNA insert clones were subsequently sequenced, and eight were found to match with the 147 bp cloned 3G1 differential display product. The longest of these, clone 18 (592 bp) was used for subsequent databank sequence homology searches.

To obtain the hBRAG cDNA 5' end, a human pre-B cell line (Nalm-6) cDNA library was employed. Specifically, ten Nalm-6 pre-B cell λZapII cDNA sublibraries size-selected for inserts ≥ 2 kb and with amplified titers of 1x10⁶ pfu/sublibrary were generated by David Kaul and kindly provided by Andre Bernards (Massachusetts General Hospital, Harvard Medical School, Boston, MA). For primary screening, 2x10⁶ PFU were plated (20,000 PFU/150 mM plate), duplicate colony lifts were performed, and the 25108-5' cDNA was used as a ³²P-radiolabeled probe for detecting hybridizing clones. Hybridizations, washes, and selection of putative clones were performed as described above. cDNA inserts were released from Bluescript via EcoRI digestions.

**DNA sequencing and sequence data analysis**

The primers for sequencing were synthesized using a Perceptive 8909 automated DNA oligosynthesizer (Perceptive Corp., Boston, MA) and purified by oligonucleotide purification cartridge chromatography (Applied Biosystems, Weterstadt, Germany). Complete DNA sequencing of the full-length hBRAG Nalm-6 cDNA clone 8-3, the EST 25108, and the C3-A11N cDNA clone 18 on both strands was performed by cycle sequencing using the Thermo Cycle Sequencing Kit (Amersham, Arlington Heights, IL) on a LICOR 4000L automated sequencer (LICOR Corp.). Sequences were analyzed, verified and assembled using the Sequencher 3.0.1
DNA analysis program (Gene Codes Corp., Ann Arbor, MI). Nucleotide and amino acid sequences were analyzed for homology or identity with known sequences in nucleotide (EMBL, GenBank, DDBJ, dbEST) and protein (Swiss-Prot, Pro-site) databases using the data base similarity search algorithms FASTA, BLAST-N and BLAST-P from the GCG software package version 8 (Madison, WI), the National Center of Biotechnology Information (NCBI) electronic mail server, and from the GenomeNet electronic mail server. Motif Searches were done using the GCG software Motifs program, and the Proscan and Blocks algorithms on the Baylor's Search Launcher server. Expressed sequence tag linkage groups were obtained through the Human cDNA database server. The predicted open reading frame was determined using the TranslateTool program and sequence alignments were done in the L-ALIGN program, both from the ExPasy server. The transmembrane domain was predicted using the TMpred program, the algorithm of which is based on the statistical analysis of TMbase, a database of naturally occurring transmembrane proteins and by the EMBL-Heidelberg program, PHDhtm. Secondary protein structure predictions were done using the Chou and Fastman algorithm for predicting β-pleated sheets and the Garnier algorithm for predicting α-helical regions, respectively, both found in the GeneWorks software package.

Genomic DNA sources and Southern blotting

Human genomic DNA was isolated from whole blood using DNAzol reagent according to the manufacturer's directions (Gibco). Baboon, Cow, Sheep, Rabbit, and Chicken genomic DNA was kindly provided by David Irwin (Toronto General Hospital, Toronto, ON). Balb/c mouse splenic/liver, whole D. melanogaster, C. elegans, and S. cerevisiae genomic DNAs were kindly provided by Juan-Carlos Zuniga Pfucker, Gabrielle Boulianne, Andrew Spence, and Brenda Andrews, respectively, all from the University of Toronto. For each digest, genomic DNA was subjected to restriction endonuclease digestion and electrophoretically fractionated at 30V on a 0.8% agarose gel in 1X TAE overnight, using a recirculator. The gel was then transferred overnight in 10X SSC to a ζ-probe® Nylon membrane (Bio-Rad, Hercules, CA), and immobilized
using a Stratalinker 2400 UV crosslinker (Stratagene). The hBRAG coding-spanning cDNA, 8-3, 5', was used as a template to generate a $^{32}$P-labeled probe by the random priming method using the Quickprime DNA labeling system (Pharmacia), and purified using Nic Spin Columns (Pharmacia). Hybridization of the probe was carried out for 24h according to directions provided by ζ-probe®. Washing was done either under low stringency (2X, 30 minutes, 5% SDS ζ-probe wash solution, 65°C), intermediate stringency (2X, 30 minutes, 5% SDS ζ-probe wash solution, 65°C; 1X, 30 minutes, 1% SDS ζ-probe wash solution, 65°C), or high stringency conditions (2X, 30 minutes, 1% SDS ζ-probe wash solution, 65°C, 2X, 30 minutes, 0.5% SDS ζ-probe wash solution, 70°C). The membranes were wrapped in plastic wrap and exposed to BioMax MS X-ray film (Kodak) for indicated times at -70°C in the presence of intensifying screens.

**In vitro transcription/translation**

The troponin T-coupled rabbit reticulocyte lysate system (ProMega, Madison WI) was used for transcription/translation reactions. Rabbit reticulocyte lysates were mixed with 1 μg of linearized pBluescript 8-3 hBRAG and T3 or T7 RNA polymerase, after which amino acid mixture without methionine plus $[^{35}S]$ methionine (Dupont) was added. After 2h at 30°C, the translated reaction mixture was resolved on a 10% SDS/PAGE gel with a rainbow protein marker (Amersham).

**Chromosomal localization of hBRAG**

For Fluorescence in situ hybridization (FISH) analysis, a 1.1 kb cDNA from the hBRAG 3' UTR (generated by BsrXI and PstI double digestion of the hBRAG 8-3 cDNA) was used to screen a P-1 derived artificial chromosome (PAC) library [388] in the Canadian Genome Analysis and Technology (CGAT) Program Physical Mapping Resource Facility. From this screening, one genomic PAC (424E18) clone was identified as positive. The regional assignment of this genomic PAC clone was determined by FISH mapping (performed at the CGAT FISH Mapping Resource Centre) of normal lymphocyte chromosomes counterstained with propidium iodide and 4',6-Diamidin-2-phenylindol-dihydrochloride (DAPI) following published methods [389]. Biotinylated
probe was prepared by nick translation and detected with avidin-fluorescein isothiocyanate (FITC), followed by biotinylated anti-avidin antibody and avidin-FITC (Oncor, Inc., Gaithersburg, MD). Images of metaphase preparations were captured by a thermoelectrically cooled charge coupled camera (Photometrics, Tucson, AZ). Separate images of DAPI banded chromosomes [390] and of FITC targeted chromosomes were obtained, pseudo colored yellow (FITC) and blue (DAPI), and merged electronically using Adobe Photoshop™ 3.0 software. The band assignment was determined by measuring the fractional chromosome length and by analyzing the banding pattern generated by the DAPI counterstained image [391] of 20 well-spread metaphases. Positive hybridization signals were visualized on both homologues in >90% of the positive spreads.

For Radiation Hybrid chromosomal mapping of hBRAG, a primer pair (hBRAG 3' UTR-S, 5' CGT GGC GTG AGA TTT GGA A 3'; hBRAG 3' UTR-A, 5' CCC CGA AGC CTG GTC TGT CT 3') was designed to amplify a 329 nucleotide product in the 3' UTR of the human hBRAG gene. A panel of 83 radiation hybrids corresponding to the Stanford Human Genome Center (SHGC) Generation 3 RH 10,000 RAD panel (Research Genetics, Huntsville, AL) was used to localize the hBRAG 3' UTR to a human genome project framework marker by PCR. PCR reactions were carried out in 25 μl reaction volumes with 25 ng of DNA, 12.5 pmol of each primer, 1.25 units of Taq polymerase (Perkin Elmer), 200 mM of each dNTP (Boehringer Mannheim), 1.0 mM MgCl₂ (Perkin Elmer), 20 mM Tris- HCl pH 8.4 and 50 mM KCl. A "hot start" was carried out with the addition of Taq polymerase and dNTP after an initial denaturation at 94°C for 5 minutes followed by 35 cycles of 94°C for 30 seconds, the optimal annealing temperature for the specific primer set for 30 seconds, 72°C extension for 30 seconds followed by a single final extension at 72°C for 15 minutes. All reactions were carried out in a DNA Thermocycler 480 (Perkin Elmer) with a mineral oil overlay. Each radiation hybrid PCR was then scored independently by two observers as positive, negative, or indeterminate. The results were then submitted to the Stanford Human Genome Center (rserver@shgc.STANFORD.EDU) for two point maximum likelihood analysis to identify linkage to SHGC framework markers in the G3
map with a LOD score of 6 or greater. Other markers in the region were then identified using the CEPH Integrated Genetic Map version v4c7 (http://www.chlc.org).

Generation and stable transfection of hBRAG expression constructs

To stably express the full length hBRAG cDNA into OCI LY8-C3P and K562 cell lines, the 8-3 cDNA was unidirectionally cloned into the BamHI and Asp178 sites of the pCEP4 eukaryotic expression plasmid (Invitrogen Corp.), a CMV promoter-based and hygromycin resistance conferring shuttle vector. The pCEP4 EBV origin of replication and the expression of EBNA-1 allow its high copy episomal replication. All ligations were performed using T4 DNA ligase (Gibco) and ligated products were electroporated into XL1-Blue E. coli. Plasmid quantities suitable for transfections were generated using pZ523™ spin columns according to the supplier's instructions (5prime->3prime, Inc., Boulder, CO). OCI LY8 C3P or K562 cells were electroporated with either empty pCEP4 vector or the pCEP4 8-3 construct using the Gene Pulser system (Bio Rad). 3x10⁷ cells in 0.75 ml of PBS were mixed with 100 μg plasmid in 100 μl of PBS, pulsed at 700 V and 25 mF and incubated at 5% CO₂ and 37°C for 48h in non-selective media. Cells were plated out at limiting dilution for 3 weeks in media with the addition of 200 μg/ml and 1 mg/ml Hygromycin (Gibco) for K562 and OCI LY8-C3P, respectively. Independent clones were expanded from plates in which <30% of wells had cell growth.
3.4 RESULTS

3.4.1 Selection of candidate cDNAs detecting RAG co-expressing transcripts

Differential constitutive RAG expression accompanies *in vitro* secondary rearrangement events in the human mature B cell line OCI LY8-C3P [223]. Whereas the parental cell line OCI LY8-C3P expresses low levels of RAG1 and RAG2 mRNA, clonally-related variants such as C3-A11N express increased levels of RAG mRNAs. As a primary screen for selecting RAG co-expressing partial gene sequences, I performed differential display RT-PCR (DD RT-PCR) on OCI LY8-C3P and C3-A11N (see experimental procedures). Using this methodology, 22 reproducible differences were found, including RAG2 itself [392]. To further select candidate RAG co-expressing cDNAs, the differential expression patterns were verified by radiolabeling the purified, reamplified cDNAs isolated from the differential display screens to probe total RNA from the original low and high RAG-expressing OCI LY8 variants. The 147 bp 3G1 amplicon was one of the reproducible differentially expressed products obtained from the DD RT-PCR screen (Figure 3.1A). When I used this amplicon as a radiolabeled probe in Northern analysis, a transcript of 5.0 kb was found to be reproducibly differentially expressed in C3-A11N (Figure 3.1B). When the 3G1 amplicon was used to probe total RNA from an array of human RAG+ and RAG- cell lines, the 5.0 kb transcript correlated strongly with RAG1 expression (see Figure 2.5, Chapter 2).

3.4.2 cDNA cloning of hBRAG

To obtain the cDNA specifically detecting the 5.0 kb transcript, the radiolabeled cDNA probe was eluted directly from the Northern blot, reamplified, subcloned into PCRII, and sequenced. Using the basic local alignment search tool (BLAST) algorithm [393], no significant sequence matches to the cloned 3G1 cDNA were found. A 592 bp cDNA insert (clone 18) was obtained by screening a C3-A11N cDNA library with this amplicon. Identical matches were found in dbEST with three 3' human fetal brain tissue-derived ESTs, all part of the tentative human consensus (THC) 114398. Additional 5' sequence was obtained by searching the TIGR (The Institute for Genomic Research)
Figure 3.1. Identification a 5.0 kb differentially expressed transcript in the high RAG-expressing variant C3-A11N by the combination of DD RT-PCR and Northern blot.

A. Representative denaturing polyacrylamide gel on which $^{35}$S dATP-labeled DD RT-PCR products from low RAG-expressing parental cell OCI LY8-C3P and high RAG-expressing C3-A11N variant are run adjacent to each other. The 3G1 differential display amplicate (denoted by an arrow) was detected using the partially-specific primer T12MG (where “M” represents the degenerate bp) and the random decamer AP3. To minimize banding from contaminating chromosomal DNA, a control for each reverse transcriptase reaction, in which no RT was added, was subjected to the same PCR reactions and represented along with the rest of the samples (-RT). Differences were considered significant only if they could be reproduced in two subclones of each variant and in duplicate PCR reactions. A sequencing ladder was also run in order to size cDNAs; the region from 150-250 bp is shown. B. Confirmation of differential expression in OCI LY8-C3P and C3-A11N. The purified 3G1 amplicate was used as a radiolabeled probe in Northern analysis of 2 µg poly(A)+ RNA from both cell variants. The 5.0 kb differentially expressed transcript is denoted by an arrow. The poly(A)+ RNAs were also probed with human RAG1 to verify differential RAG expression and with β-actin cDNAs to adjust for loading variation. Exposure times for β-actin, RAG1, and 3G1 are 6, 24, and 72 h, respectively. Positions of the 0.24-9.49 kb RNA ladder are indicated on the far left.
human cDNA database (HCD) for other linked ESTs. Two non-overlapping ESTs from a separate THC were found to be contained within the same clone (human cDNA clone 25108, a product of the IMAGE consortium sequencing effort) as THC 114398, indicating that these linked ESTs represented the 5' end of 25108. After complete sequencing of 25108 (purchased from Research Genetics, Huntsville, AL), re-entry of the entire 25108 sequence into BLAST revealed no further matches with any other nucleotide sequences and no potential open reading frames were found. A 1.5 kb fragment representing the 5' end of 25108 (25108-5') was then used to screen a human Nalm-6 pre-B cell library and the four largest unique clones were selected for sequence analysis (Figure 3.2). Consistent with restriction mapping analysis and a single transcript seen in Northern blot analysis of Nalm-6 total RNA, these clones were found to have identical overlapping sequences and are therefore derived from the same cDNA. The relative positions of the original 3G1 differential display product, clone 18 from the C3-A11N cDNA library, cDNA clone 25108 (containing hBRAG ESTs), and the various isolated Nalm-6 cDNA clones are shown in Figure 3.3.

3.4.3 Analysis of hBRAG nucleotide and predicted amino acid sequence

The complete hBRAG cDNA (4382 bp) contains a putative open reading frame encompassing nucleotides 207-1715, predicting an encoded protein of 503 aa (Figure 3.4A). The ATG at nucleotides 206-208 has features of the genuine initiation methionine; it is embedded in a near-consensus Kozak sequence [394], and there are nearby upstream stop codons in all three reading frames. The hBRAG nucleotide sequence is also composed of 2666 bp of 3' UTR sequence containing a classical AATAAA polyadenylation signal, and a poly(A) tail. Other clones analyzed in the coding region sequence were found to be identical with one exception, the presence of an in-frame 12 base pair deletion at nucleotides 1035-1046 in cDNA clones 5-1 and 6-6. Since this deletion is present in more than one clone, it is likely to represent an alternative splicing variant rather than a cloning artifact.
Figure 3.2 Isolation of the complete hBRAG cDNA 8-3 and other related clones from a human Nalm-6 pre-B cell cDNA library

Nalm-6 human pre-B cell cDNA sublibraries were screened using the EST 25108-5' (corresponding to hBRAG 3' UTR region) as a probe. Shown is a 1% EtBr agarose gel of EcoRI digests of in vivo excised inserts from 30 screens. Sizes of these inserts are indicated below. Position markers (in bp) of a 1 Kb DNA ladder is shown at the far left.
Figure 3.3. Map of hBRAG cDNA cloning strategy.

Denoted are the positions of the differential display amplicon 3G1, the C3-A11N cDNA library clone #18, and the human ESTs representing the hBRAG 3' UTR. A probe derived from the 5' end of EST clone 25108 (25108-5') was used to identify Nalm-6 cDNA clones. The four largest cDNA clones obtained from this screen are shown. The largest clone (clone 8-3, 4.2 kb) was sequenced in its entirety on both strands. The other clones were mapped relative to 8-3 by restriction mapping and partial sequencing (dashes in clones 4-7, 5-1, and 6-6 denote regions not sequenced). An EST corresponding to the 5' end of the hBRAG cDNA is also shown. The 25108-5' probe used to screen for the 5' end of the hBRAG cDNA is denoted as a stippled line. A restriction map of the resulting complete hBRAG cDNA is also shown. The filled area represents the hBRAG open reading frame, and the thick lines indicate the 5' and 3' UTRs. A scale bar (in kb) is shown at the top.
EST 00496 (381 bp)

Nalm-6 λZapII clone 8-3 (4.23 kb)

Nalm-6 λZapII clone 6-6 (3.42 kb)

Nalm-6 λZapII clone 4-7 (3.27 kb bp)

Nalm-6 λZapII clone 5-1 (2.93 kb)

EST T80458 (443 bp)

EST H17845 (497 bp) EST R39021 (367 bp)

HCD linkage clone 25108 (2.64 kb)

C3-A11N λZapExpress clone 18 (592 bp)

3G1 D.D. ampiclon (157 bp)
Figure 3.4. Nucleotide sequence of the hBRAG cDNA, predicted amino acid sequence, and secondary structure of the encoded protein.

A. hBRAG cDNA nucleotide sequence. The sequence shown (EMBL accession number AF026477) is obtained from cDNA clone 8-3 and an additional overlapping 144 bp of 5' sequence derived from human EST #00496 (381 bp). The numbers to the right of each row refer to the nucleotide position (upper sequence) or to the aa position (lower sequence) in the predicted hBRAG protein. The longest open reading frame, encompassing nucleotides 207-1715, predicts a protein of 503 aa. The initiation methionine is in bold and the stop codon is denoted by an asterisk. The polyadenylation signal AATAAA close to the polyadenylated 3' end of the cDNA is shaded. The hydrophobic membrane-spanning region of the protein and the potential rapid degradation signals in the 3' UTR of the cDNA are underlined. Potential sites for N-glycosylation [N-X-(S/T)] are boxed. Cysteine residues that represent potential palmitylation sites are circled. The original differential display amplicon 3G1 is double underlined. Other features of the cDNA sequence include a phosphorylation site at nucleotides 316-324 (aa 38-41) for protein kinase C [(S/T)X(R/K)] and at nucleotides 441-452 (aa 78-81) for the cyclic AMP-dependent protein kinase [(R/K)(R/K)X(S/T)] and an in-frame intervening sequence at nucleotides 1035-1046 observed in clone 6-6 but not in clone 8-3 representing either a possible splice junction or cloning artifact. B. Hydropathy plot of the putative hBRAG protein using the hydrophobicity indices of Kyte and Doolittle [395]. Potential transmembrane regions are delineated by a cutoff line at 160. A schematic diagram of the predicted domains of the hBRAG protein including four putative N-glycosylation sites in the extracellular domain is also shown. C. SDS/PAGE gel analysis of the in vitro translated hBRAG gene product. Linearized pBluescript 8-3 hBRAG was transcribed in vitro, using either the T3 promoter, to generate a sense RNA or the T7 promoter to generate an anti-sense RNA. The two RNAs were translated in vitro in the presence of 35S-methionine. Also shown is a negative control reaction (-), in which no cDNA was added, and a positive control reaction (+), in which RNAs from a transcribed MLk-3 kinase construct was translated into the expected 98 kDa product.
Nucleotides 453-503 (aa 82-98) encode a highly hydrophobic sequence composed of 17 aa characteristic of a transmembrane segment, as predicted by a Kyte and Doolittle hydropathy plot (Figure 3.4B) [395]. A second potential transmembrane domain located near the C-terminus (aa 458-475) also met the hydrophobicity criteria, but probably does not constitute a genuine transmembrane region based on the consecutive highly charged residues that are also present in this sequence. The predicted overall structure of the hBRAG protein conforms to a type II transmembrane glycoprotein: the sequence has no discernible signal peptide, there is a highly positive-charged cytosolic region directly flanking the putative transmembrane corresponding towards the C-terminal end of the protein, and the C-terminal extracellular region of 405 aa contains four potential N-glycosylation sites. The 81 aa cytoplasmic tail contains putative palmitylation sites, and PKC and cAMP-dependent protein kinase phosphorylation sites. BLASTP searches reveal no significant homology to any other known protein. Secondary structure considerations [396] and alignment of the hBRAG extracellular domain with domains present in other type II transmembrane molecules, revealed no recognizable regions that would allow placement into a particular family, such as signal-transducing proteins of the Ig superfamily [397], the C-type lectin supergene family [398], or the TNF-α receptor family [340]. To confirm that the authentic open reading frame was correctly predicted, the full length 8-3 cDNA clone was transcribed and translated in vitro. The labeled protein product was analyzed on a SDS-PAGE gel, as shown in Figure 3.4C. A doublet lies between the 69 kDa and 49 kDa markers, the major product whose size is consistent with the predicted non-glycosylated size of the hBRAG protein. It is presently unknown if this doublet represents related proteolytic products, or different isoforms generated from alternative translational initiation sites.

3.4.4 Expression pattern of hBRAG cDNA

To investigate the message size(s) and expression pattern of hBRAG, I employed two separate probes representing different regions of the hBRAG cDNA: 8-3, 5', spanning the majority of the hBRAG coding sequence, and 25108-5', spanning a large portion of the hBRAG cDNA 3' UTR
(Figure 3.5A). With either of these two probes, a single 5.0 kb transcript was detected only in human cell lines that were of the B cell lineage and that expressed the 6.6 kb RAG-1 transcript (Figure 3.5B). These results were similar to those previously observed with the 3G1 differential display amplicon. Either low or undetectable levels of hBRAG message were observed in total RNA from various non-lymphoid lines (Figure 3.5C). The ~3 kb ubiquitously-expressed transcript seen previously with the radiolabeled 3G1 amplicon (Figure 3.1B) was no longer detected, suggesting this smaller transcript represented a cross-hybridizing, non-specific product picked up only with the original differential display product (likely because of its short length and/or its bias towards repetitive sequence). Thus, hBRAG is expressed at the RNA level only in B cells that express RAG and not in human T cell lines, irrespective of their expressed levels of RAG1 transcripts (Figure 3.6B).

The distribution of hBRAG mRNA in a wide array of human tissues was determined by RNA dot blot analysis (Figures 3.6A-B). hBRAG is expressed at low levels in bone marrow tissue, which may be consistent with the Northern analysis of B cell RAG+ lines, since unfractionated bone marrow only contains a small percentage of RAG-expressing B cells ([62] and Figure 3.6C). Surprisingly, hBRAG is expressed at highest levels in secondary lymphoid tissues that are highly B cell-enriched such as fetal spleen, adult spleen, peripheral blood leukocytes (PBL), and lymph node. hBRAG is also expressed in heart, ovary, stomach, and brain putamen, and at low but detectable levels in most other tissue types (Figure 3.6C). hBRAG is not expressed in thymus, consistent with the Northern analysis of RAG+ T cell lines. Because these tissues are enriched for RAG-expressing thymic precursors, it is clear that hBRAG expression is B cell and not T cell-associated (Figure 3.6C). Consistent with RNA dot blot results, Northern analysis of human tissues demonstrate that hBRAG is expressed in bone marrow and tonsil, but not in thymus or placenta (Figure 3.6D). Furthermore, as with Northern analysis of RAG+ B cell lines, a single 5.0 kb transcript is observed in these B cell-enriched tissues.
Figure 3.5. Co-expression of hBRAG and RAG1 message in human early B cell lines.

A. Location of probes in the complete hBRAG cDNA used for expression analysis. Probe 8-3, 5' (nucleotides 145-1899) spans 63 bp of 5' UTR sequence, the entire coding sequence (denoted by a boxed region), and 184 bp of 3'UTR. Probe 25108-5' (nucleotides 1899-3406) spans a large portion of the 3' UTR. The position of the differential display amplicon 3G1, used in preliminary expression analysis, is also shown. B. Northern blot analysis of 15 µg total RNA for hBRAG message in various human B and T cell lines representing the different developmental stages shown, using 25108-5' and 8-3, 5' as radiolabeled probes. The same blot was then sequentially hybridized with human RAG1 and β-actin cDNA probes. Exposure times for 25108-5', 8-3, 5', RAG1, and β-actin are 14, 12, 8, and 6 h, respectively. C. Northern blot analysis of 20 µg total RNA from various non-lymphoid human cell lines was performed as in B. The positive controls for hBRAG expression are 697, C3-A11N, and REH. Exposure times for 8-3, 5', RAG1, and β-actin are 36, 24 and 8 h, respectively.
Figure 3.6. B-lymphocyte lineage specificity of hBRAG expression in human tissues.

A. Type and position of human poly(A)+ mRNAs on a normalized human Master blot (CLONTECH). B. Hybridization of radiolabeled 8-3, 5' hBRAG coding region probe to a human RNA Master blot (CLONTECH). C. Hybridization of radiolabeled RAG1 coding region probe to human RNA Master blot (CLONTECH). Exposure times for 8-3, 5' and RAG1 are 36 and 24 h, respectively. D. Northern blot analysis of 2 μg human poly(A)* mRNAs from primary and secondary lymphoid tissues sequentially hybridized with 8-3, 5' and β-actin. Exposure times were 24 h for 8-3, 5' and 6 h for β-actin.
### A

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>whole brain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>occipital lobe</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>heart</td>
<td>aorta</td>
<td>skeletal muscle</td>
<td>colon</td>
<td>bladder</td>
<td>uterus</td>
<td>prostate</td>
<td>stomach</td>
</tr>
<tr>
<td>d</td>
<td>testis</td>
<td>ovary</td>
<td>pancreas</td>
<td>pituitary gland</td>
<td>adrenal gland</td>
<td>thyroid gland</td>
<td>salivary gland</td>
<td>mammary gland</td>
</tr>
<tr>
<td>e</td>
<td>kidney</td>
<td>liver</td>
<td>small intestine</td>
<td>spleen</td>
<td>thymus</td>
<td>peripheral leukocytes</td>
<td>lymph node</td>
<td>bone marrow</td>
</tr>
<tr>
<td>f</td>
<td>appendix</td>
<td>lung</td>
<td>uterus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>fetal brain</td>
<td>fetal heart</td>
<td>fetal kidney</td>
<td>fetal liver</td>
<td>fetal spleen</td>
<td>fetal thymus</td>
<td>fetal lung</td>
<td></td>
</tr>
</tbody>
</table>

### B

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### C

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### D

<table>
<thead>
<tr>
<th>Heart</th>
<th>Placenta</th>
<th>Thymus</th>
<th>Bone Marrow</th>
<th>Tonsil</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0 kb</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 kb</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**hBRAG (8-3, 5')**

**β-actin**
Because RAG1 total RNA can be upregulated by sIg crosslinking (anti-μ ligation) or by treatment with second messengers in the OCI LY8 line and its variants [75], I was interested in determining if hBRAG is co-regulated with RAG1 upon such induction. I found that upon either anti-μ treatment or by stimulation with second messengers, hBRAG transcript levels were not altered (data not shown).

3.4.5 Southern analysis and chromosomal localization of the hBRAG gene

To begin to characterize the genomic organization of hBRAG, I carried out Southern blotting using the hBRAG coding region probe. Human genomic DNA was digested with several restriction endonucleases, electrophoresed, transferred, hybridized to 8-3, 5', and washed under high and low stringency conditions (Figure 3.7A). Under high stringency wash conditions, one ~ 13 kb predominant band and a smaller one of 3.8 kb is seen in the BamHI digest. The EcoRI genomic digest shows a single predominant band at 8.5 kb. The HindIII genomic digest has predominant bands at 4.9, 4.7, and 4.1 kb (total of 13.8 kb). Two bands, of 4.6 and 2.2 kb in length, obtained with the PstI digest, represent a total of 6.8 kb. Under low stringency conditions, the same predominant bands plus some additional, low-signal intensity bands were seen. The fact that only one or two predominant bands hybridized in all four digests (even under low stringency conditions) suggests that the hBRAG gene is likely not a member of a multigene family. Moreover, the number and summed sizes of bands present in each digests indicates that the hBRAG gene is likely to be present as a single copy gene in the human genome with most of the coding information present within as little as 6.8 kb. The presence of up to three predominant bands in the digests may be due to the presence of restriction site(s) within one or more introns, since the 8-3, 5' fragment does not contain an internal EcoRI site.

To study whether hBRAG is conserved across species, I digested genomic DNA from various eukaryotic species with PstI, and probed the DNAs with 8-3, 5'. I found that under low stringency washing conditions, this gene is highly conserved across eukaryotic species, as
Figure 3.7. Genomic Southern blot analysis of hBRAG.

A. Southern blot analysis of human genomic DNA digested with various restriction enzymes. 10 µg of human DNA was digested with BamHI, EcoRI, HindIII, and PstI, hybridized to the 8-3, 5' hBRAG coding region fragment, and washed under high stringency conditions (2X, 30 minutes, 1% SDS Zeta-probe wash solution, 65°C; 2X, 30 minutes; 0.5% SDS Zeta-probe wash solution, 70°C) or under low stringency conditions (2X, 30 minutes, 5% SDS wash solution, 65°C). Exposure time for both conditions was 7 days. B. Southern “Zoo blot” analysis of genomic DNA digested with EcoRI. Genomic DNA (10 µg) from various species was digested with EcoRI, probed with 8-3, 5', and washed under the same high or low stringency conditions as described for A, or under intermediate stringency conditions (2X, 30 minutes, 5% SDS Zeta-probe wash solution, 65°C; 1X, 30 minutes, 1% SDS Zeta-probe wash solution, 65°C). Exposure time for the blot washed under high stringency was 21 days, for the intermediate and low-stringency blots, 14 days. C. Southern “Zoo blot” analysis of genomic DNA digested with PstI. Genomic DNA (25 µg) from various species was digested with PstI, probed with 8-3, 5', and washed under the same high, intermediate, or low stringency conditions as described for B. For all blots, positions of size markers (in bp) are indicated at the far left.
predominant bands were present in all species from human to yeast (Figure 3.7C). Under intermediate stringency conditions, predominant bands appear present only in the vertebrate species, suggesting a particularly strong degree of conservation in these species. Under high stringency conditions, the predominant bands of 4.6 kb and 2.2 kb are only detected in human and baboon DNA, two species with ~95% genomic DNA identity. The same type of Southern "zoo blot" analysis was performed using genomic DNA restricted with EcoRI, and again conservation was seen across all species under low stringency conditions, and a stronger conservation was seen in vertebrate species under intermediate stringency conditions (Figure 3.7B).

To localize the hBRAG gene, an hBRAG genomic PAC clone identified by the ACGT corporation was used in FISH analysis (described in Materials and Methods). hBRAG maps to human chromosome 10q26 (Figure 3.8), which is part of a region syntenic to mouse chromosome 19 (MGD cM position 45-51.0). A more precise physical map of this region was generated through the use of the somatic cell hybridization technique, which assigned hBRAG to the GDB region D10S244E of human chromosome 10. An analysis of physical and genetic maps of this region does not reveal any mapped common or rare human immunologic diseases.

3.4.6 A role for the hBRAG gene product in the regulation of RAG1

To examine the possibility that the hBRAG gene product may regulate the expression of RAG1 transcripts, hBRAG was overexpressed by stable transfection of the 8-3 cDNA in pCEP4 (described in Materials and Methods) in either the non-lymphoid cell line K562 (non RAG-expressing), or the OCI LY8 B cell parental clone, OCI LY8-C3P, (expressing low levels of RAG1 transcripts) (Figure 3.9). No induction of RAG1 mRNA expression was observed in transfected K562 clones expressing the hBRAG cDNA, but relative increases in RAG1 mRNA (approximately four to eight fold as compared to untransfected OCI LY8-C3P cells) were seen in four of five independent OCI LY8-C3P clones. Although it is unclear why one of the transfectants does not demonstrate increased RAG1 mRNA, it is possible that this particular transfectant has
Figure 3.8. Chromosomal localization of the hBRAG gene.

FISH analysis with hBRAG genomic PAC clone generated signals on both chromatids (denoted by arrows) of chromosome 10 at band q26. The region is shown in more detail next to a banding ideogram.
Figure 3.9. Increased levels of RAG-1 mRNA in OCI LY8 cell lines transfected with the full length hBRAG cDNA.

Shown is Northern hybridization of 15 μg of total RNA from untransfected C3-A11N, K562, and OCI LY8-C3P cells, and from independent K562 and OCI LY8-C3P clones stably transfected with empty pCEP4 vector (designated K562 pCEP4-8,11,12, 14 and OCI LY8-C3P pCEP4-2,4, respectively) or with pCEP4 8-3 constructs (designated K562 8-3 pCEP4-2, 3, 4, 5, 6, 8 or OCI LY8-C3P 8-3 pCEP4 1, 2, 3, 4, 6, respectively). After hybridization to the radiolabeled hBRAG coding region probe, the same blot was then sequentially stripped and re-hybridized to human RAG1 and β-actin cDNA probes. Exposure times for hBRAG, RAG1, and β-actin are 48, 24, and 8 h, respectively. RAG1 mRNA signal intensities in the OCI LY8-C3P clones transfected with either the empty vector or the pCEP4 8-3 construct are shown below the blot in terms of expression relative to the untransfected control RAG1/β-actin signal ratios, which has been arbitrarily set at 1. Note that the C3-A11N control lane and the rest of the K562 transfectant lanes are from different portions of the same gel and have been fused together. The OCI LY8-C3P transfectants were run on a separate gel.
undergone a mutation event either in the RAG1 gene or in one of the products that promote its expression. In this context, it indicates that potentially the link between hBRAG and RAG1 is indirect, and may be mediated by other proteins that fluctuate or are mutant in independent clonal derivatives. Alternatively, this clone may express hBRAG at such high levels that it downregulates RAG1 expression. Overall, these data suggest that hBRAG may be involved in enhancing human RAG1 mRNA transcription and/or stabilization when introduced in cells that already express low levels of RAG mRNAs and/or in the context of the B cell lineage.
3.5 DISCUSSION

I have described a novel human cDNA initially isolated through use of DD RT-PCR of low and high RAG-expressing variants of the human B cell line OCI LY8-C3P. I have termed this gene hBRAG, or human B-cell RAG-associated gene. hBRAG encodes a 503 aa novel type II transmembrane glycoprotein that is conserved across species. It is expressed as a single, 5.0 kb transcript, detectable at highest levels in human tissues enriched in B cell progenitors and RAG-expressing B cell lines, and is not found in T-cell lines or thymus tissue.

Typically, type II (inverted membrane orientation) transmembrane proteins are ubiquitously expressed, function as enzymes, and are associated with the golgi-membrane. Many of these are involved in post-translational modification reactions [399]. However, of the cloned type II glycoproteins identified as having important immune functions, many are plasma-membrane associated receptors belonging to families with defined domains and are expressed in a tissue-specific manner. One category includes certain members of the TNF/nerve growth receptor family, a family characterized by sharing restricted homology in the C-terminus i.e. extracellular region [341]. These include the ligands for CD27, CD30, and CD40 (all of which are involved in the co-stimulation of B-T cell interactions), the ligand for FAS, and the TNFα and β receptor components. A second category is the LY49 receptor family, all of which have C lectin-like recognition domains (CRD) [343]. These include the murine NK cell markers LY49 and NKPR1, the low-affinity IgE receptor (CD23), the early T and B cell activation marker CD69, the IL-1 receptor α and β chains, and the early-B cell differentiation marker Lyb-2. Finally, the type II molecule BP-1 belongs to the zinc-dependent metallopeptidase family, is expressed in murine early B cell subsets, and may function as an amino-peptidase enzyme rather than a receptor [348].

Computer searches against all databases revealed no significant match to hBRAG, not even among molecules with inverted membrane orientation. Furthermore, I was not able to find regions of homology in hBRAG to any of the domains characterizing the various aforementioned type II
protein families or other receptor family domains such as Ig-like domains, leucine-rich repeats (LRR), or Selectin-like domains. Southern blot analysis would seem to reinforce the possibility that hBRAG does not belong to a known multi-gene family since a coding-region probe fails to detect multiple predominant bands under high or low stringency (Figure 3.7A, B). hBRAG does show some limited identity in a C-terminal region of the extracellular domain of heparan N-sulfotransferase (21.9% over a 199 aa stretch), a golgi-membrane associated, ubiquitously expressed type II transmembrane glycoprotein involved in controlling the extent of polysaccharide chain modification through catalysing the N-sulfonation of heparan sulfate [400]. Because short transmembrane regions are often indicators of golgi-association [401], the 17 residue transmembrane region of hBRAG may be an indication that it is a golgi-associated protein rather than a transmembrane receptor. Experiments with hBRAG anti-peptide antibodies to determine the subcellular location and expression of the hBRAG protein are described in Chapter 4.

The hBRAG gene maps to the long arm of chromosome 10 (10q26), as determined by FISH analysis. The chromosomal region 10q24-q26 is part of a known segment that is syntenic between human chromosome 10 and mouse chromosome 19 [402]. In this region are genes encoding various enzymes including lipase A, glutamic oxaloacetic transaminase, phosphoglyceromutase 1, and terminal deoxynucleotidyl transferase (TdT) as well as several multi-spanning membrane adrenergic and serotonin receptors [403]. Assignment of the hBRAG gene to human chromosome 10q26 adds another gene to this conserved syntenic region. Of significance, the type II transmembrane glycoprotein heparan N-sulfotransferase, with which hBRAG shares limited identity, also maps to this region, suggesting a possible functional and evolutionary relationship between these two genes. Radiation hybrid mapping has further localized the hBRAG gene to the GDB assigned region D10S244E of human chromosome 10. The genes most closely linked to hBRAG physically (within the 100-200 centiRay, where 1 centiRay is ~3.7 megabases; corresponding to the cytological region 10q26.2-10q26.3) are: the GTP-binding protein, fibroblast growth factor receptor 2, and the CoA hydratase genes [404]. Although this region of human
chromosome 10 is not associated with any immunological diseases, deletion and translocation events in this region have been linked to various non-lymphoid abnormalities [405-407]. Since hBRAG is expressed at low but detectable levels in several non-lymphoid tissues (Figure 3.5B), it remains to be determined if this gene has any involvement in such non-lymphoid abnormalities.

Although my initial expression analysis indicated that hBRAG is detected in human pro and pre-B cell lines co-ordinately with RAG1 expression, I subsequently observed substantial amounts of hBRAG expression in B cell-enriched secondary lymphoid tissues in addition to bone marrow. The expression of hBRAG in PBL, lymph node and spleen may be due either to subsets of RAG-expressing B cells in secondary lymphoid tissues, or to expression of hBRAG mRNAs in the absence of RAG1 at certain developmental stages. Consistent with the former possibility are recent reports of re-expression of RAG transcripts in germinal center tissues [71-73] and Figure 3.4C demonstrates low but detectable levels of RAG expression in secondary tissues including fetal and adult spleen, adult lymph nodes and PBL. Further analysis of purified pre-B and B cell subsets by RT-PCR is in progress to ascertain how hBRAG and RAG expression correlate in these tissues.

The expression pattern of hBRAG is striking in that it is expressed in RAG-expressing B cell lines and tissues but not in RAG-expressing T cell lines or in adult or fetal thymus. This suggests that hBRAG may be a useful diagnostic marker in human B cell ontogeny and may play an important role in B cell development. It is also noteworthy that my Southern analysis of this gene indicates a high degree of evolutionary conservation across vertebrate species. The apparent conservation and uniqueness of this gene suggests that it may interact with conserved substrates. In this context, the RAGs themselves, which act on evolutionarily conserved recombination signal sequences, are also conserved throughout vertebrate evolution [37].

It has been shown that RAG transcripts can be modulated both positively, and negatively, by the manipulation of various signaling pathways, such as Ig or TCR ligation [40, 62, 75, 234] or various signaling second messengers, such as cAMP, Protein Kinase C, and protein phosphatases.
l and 2A ([219, 254]; reviewed in Chapter 1). Experiments in which I have stably transfected the hBRAG cDNA into the low RAG-expressing mature B cell line OCI LY8-C3P and the non-lymphoid line K562 suggest that, at least in the context of the B cell lineage, the hBRAG protein may be involved in inducing transcription and/or stabilization of RAG1 mRNA transcripts. Because it is presently unknown whether the hBRAG protein functions as an enzyme or a receptor, this product could either be a component of a B-cell specific signaling pathway or, alternatively, may interact with (or post-translationally modify) other B-cell specific components that regulate RAG mRNA transcripts. One attractive candidate downstream factor that hBRAG potentially may be regulating are the E2A gene products, since mice with the E2A^- phenotype lack RAG1 mRNA transcripts in B cells, but not in T cells [207]. Future experiments will be aimed at more precisely determining the molecular mechanisms implicated in hBRAG-induced alterations in RAG expression and the lineage specificity of this regulation. Other studies will address the molecular associations of hBRAG in addition to its potential functions in B lymphocyte differentiation and antigen receptor recombination.
CHAPTER 4:

Biochemical characterization of hBRAG as a potential BCR signal-enhancing glycoprotein dimer that associates with phosphorylated proteins in resting B cells
Acknowledgements

Contents of this Chapter have been submitted for publication under the title *Biochemical characterization of hBRAG as a potential BCR signal-enhancing glycoprotein dimer that associates with phosphorylated proteins in resting B cells*, by Laurent K. Verkoczy, Barbara-Anne-Guinn, and Neil L. Berinstein. Barbara Guinn made biotinylated versions of hBRAG antibodies and performed all flow cytometry-based experiments involving intracellular and cell surface staining. The Synpep Corporation generated the hBRAG peptides and performed the rabbit immunizations. All other work was performed by myself. I wish to thank members of the Reithmeier lab for technical help, in particular Jeff Charuk, Milka Popov, and Jani Quilty, for their expertise and valuable advice in affinity purification, cell-free translation, and cell-surface biotinylation assays, respectively. This work was supported by the National Cancer Institute of Canada (NCIC grant # 7286).
4.1 ABSTRACT

To begin biochemical characterization of the hBRAG protein, I generated affinity-purified polyclonal antibodies against N and C-terminal epitopes of the hBRAG protein. The binding specificity of these antibodies was characterized in various contexts; against BSA-conjugated peptides, whole cell lysates of endogenously expressed hBRAG+ cell lines, against hBRAG proteins expressed in stably transfected K562 cells, and against cell-free synthesized products. Immunoblotting and immunoprecipitation experiments with these antibodies demonstrated that the antibody reagents detected multiple isoforms of hBRAG. Taken together, the results suggest that hBRAG can be expressed in B cells as a membrane-integrated glycoprotein disulfide-linked dimer. This feature is consistent with earlier structural predictions of a TM region with 4 N-glycosylation sites in the hBRAG extracellular domain. A dimeric structure is also consistent with the ability of almost all type II lymphocyte receptors to form higher order structures. Flow cytometry and cell surface biotinylation experiments revealed that hBRAG subcellular localization includes both a cell surface and intracellular component. Immunoprecipitation with hBRAG antisera detected co-immunoprecipitated phosphorylated proteins in resting B cells, including the protein tyrosine kinase Hck, which may be subsequently de-phosphorylated upon BCR ligation. Consistent with its cell surface expression and possible link to BCR signaling, experiments in which α-hBRAG antibodies were used to generate early activation signals suggest a modest, but specific element of tyrosine phosphorylation occurring through a putative hBRAG receptor. These experiments also suggest that hBRAG may also be involved in positively enhancing BCR ligation–mediated early activation events. Overall, these results are consistent with a function for hBRAG as a cell-surface signaling receptor molecule. Coupled with the earlier observation that hBRAG expression correlates with early and late B cell-specific RAG expression, this also suggests that hBRAG may mediate potentially important regulatory signals in B cell signaling pathway that are important for B cell development and/or regulating B-cell-specific RAG expression.
4.2 INTRODUCTION

In general, type II transmembrane proteins function as intracellular enzymes involved in biosynthesis and post-translational modification pathways [339]. However, amongst those with immunological function, there are only two examples of type II TM molecules that function as enzymes, the intracellular invariant chain (reviewed in [358]) and the cell-surface expressed BP-1 neutral endopeptidase (reviewed in [139]). The rest, which include the LY49 lectin family and the TNF receptor family, function exclusively as cell-surface signaling molecules. The hBRAG protein is a novel type II TM molecule, which appears to have a role in B lymphocyte development and RAG regulation in this lineage, as suggested by expression and transfection analysis. However, this protein has no homology to any known protein, making it difficult to ascertain clues as to its modus operandi in the aforementioned functions. To aid in elucidating hBRAG function, it would be informative to biochemically characterize this molecule.

There are at least two distinct ways hBRAG could function in the processes of B cell development and RAG regulation. One would be indirectly, as an intracellular enzyme involved in transport, synthesis, or post-translational modification of B cell-specific factors. The other would be directly, as a B cell signal-transducing molecule. In this context, the generation of reagents against hBRAG would be useful not only in the biochemical characterization of this protein, but would allow the analysis of the cellular distribution of this molecule and the identification of hBRAG interactions with other known proteins.

In this study, I generate antibodies to hBRAG and characterize their specificities in various contexts. This has allowed me to extend structural predictions that hBRAG is a glycosylated, membrane integral, disulfide-linked dimer, that may associate with other phosphorylated proteins in resting, but not BCR cross-linked B cells. I also provide some preliminary evidence that suggest an hBRAG function as a cell surface receptor capable of transducing signals alone or possibly in conjunction with the BCR. Overall, these data suggest that, like other lymphocyte-specific type II molecules, hBRAG may function in developmental and lineage-specific signaling.
Furthermore, like other receptors, including IL-7, CD19, and the BCR, this protein may function as a signaling-mediating regulator of RAG expression in B cells. Finally, our finding that an intracellular fraction of hBRAG exists cannot exclude a separate, additional role for hBRAG as an intracellular enzyme or chaperone.
4.3 MATERIALS AND METHODS

Cell lines, human tissues, and tissue culture

The cell line panel used for peptide competition assays included the following human cell lines: REH (pro-B), 697 (pre-B), A8-6P (mature B), HeLa (fibroblast), U937 (pro-monocytic). These have been described in detail previously [49]. All human cell lines (unless otherwise indicated) were cultured in RPMI 1640 media (Wisent, Que, Canada) supplemented with 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, and 15% bovine calf serum (Wisent). Mouse cell lines were additionally supplemented with 50 µM 2-ME. All cells for all experiments were incubated at 37°C and 5% CO₂ and harvested in the log growth phase. Tissue protein medleys of human heart, placenta, and bone marrow were obtained from Clontech (Palo Alto, CA), human thymus tissue was kindly provided by David Hogg (University of Toronto, Toronto, ON), and PBL were obtained from the Sunnybrook Health Sciences Centre.

For extracting purified human mature B cells, juvenile tonsils were recovered post tonsillectomy. Tonsillar tissue was homogenized manually and subsequently strained through a 70 µm nylon cell strainer (Falcon, Franklin Lakes, New Jersey). The strained cells in the flow-through were then centrifuged in Ficoll-Paque (Pharmacia Biotech, Sweden) density gradient. Isolated lymphocytes were washed in RPMI 1640 and B cells were isolated by negative selection using neuraminidase treated sheep red blood cells (Cederlane, ON) to rosette T lymphocytes, as per the manufacturer's protocol. All B cell-enriched preparations were analyzed by flowcytometry for percent B lymphocyte purity based on CD19⁺ expression.

hBRAG peptide synthesis, polyclonal antibody production and affinity purification

Peptides derived from the most hydrophilic N and C-terminal regions of the cloned hBRAG protein, consisting of residues 82-98 and 479-503, respectively, were synthesized by Synpep Corporation (Dublin, CA). The peptides were constructed on a Pharmacia Biolynx Automated Peptide Synthesizer and were purified by reversed-phase HPLC. For immunizations, the synthesized peptides were conjugated to Keyhole Limpet Hemocyanin (KLH) with
gluteraldehyde. New Zealand White rabbits (two animals per peptide conjugate) were immunized by initial intramuscular injections of 500 µg of the peptide-KLH conjugate emulsified in complete Freund’s Adjuvant. Rabbits were boosted with three subsequent subcutaneous injections of 250 µg of the antigen in Freund’s incomplete adjuvant every 2 weeks before cardiac puncture. Rabbits were bled one week after each boost for serum collection. For affinity purifications of polyclonal antibodies from serum bleeds, synthesized peptides were conjugated to BSA with gluteraldehyde and peptide-conjugated affinity chromatography columns were prepared by coupling the BSA-conjugated peptides to a CNBr-activated Sepharose 4B bead matrix (Pharmacia), according to the supplier’s instructions. 1.5 ml sera/column was loaded on the column, after extensive washing of the columns in 1X TBS, α-hBRAG peptide antibody fractions were eluted in Pierce elution buffer (Pierce, Rockford, IL) and stored at 4°C.

Pulse and steady-state labelings and immunoprecipitations

For pulse labeling experiments, 1x10^7 A8-6P cells were washed in 1X PBS and resuspended in a 10 ml volume of either L-Methionine-free RPMI media, and incubated with 10 μCi/ml Easytag L-[^3S]-Methionine (Dupont NEN, Boston, MA) for 3 h. Alternatively, for steady state labeling experiments, 1x10^7 A8-6P cells were resuspended in 10 ml of normal RPMI media, and incubated with 10 μCi/ml Easytag L-[^3S]-Methionine (Dupont NEN) for 24 h. Cell pellets were solubilized in 6 ml of radioimmunoprecipitation assay (RIPA) buffer (1% deoxycholate, 1% Triton-X-100, 0.1% SDS, 0.15M NaCl, 1 mM EDTA, and 10 mM Tris-HCl, pH 7.5) supplemented with the contents of a complete protease inhibitor mixture tablet (Boehringer Mannheim Corp., Indianapolis, IN). Nuclei, cytoskeletal components, and unlysed cells were removed by centrifugation at 4°C for 10 min at 14,000 g. For each immunoprecipitation reaction, lysates were precleared before immunoprecipitation by incubating 1 ml of cell lysate with 40 µl of a 1:1 suspension of Protein A Sepharose (Pharmacia) for 1 h at 4°C. Lysates were then incubated with gentle mixing overnight at 4°C with pre-immune sera, crude N-terminal antisera, or affinity purified rabbit IgG control, or N-terminal α-hBRAG antibodies (antibody concentrations used are detailed in the figure legends). 20 µl of a 1:1 suspension of Protein A
Sepharose: 1X RIPA buffer was added, and samples mixed for an additional 2 h at 4°C. Resins were washed twice with 1 ml aliquots of ice cold of the relevant immunoprecipitation buffer, and immunoprecipitates were solubilized in 20 μl of 2X electrophoresis sample buffer (New England BioLabs Inc., Beverly, MA) supplemented with 1% 2-ME and 0.042 mM DTT.

**In vitro transcription/translation and co-translational processing studies**

Transcription reactions were carried out using 0.5 μg of linearized pBluescript 8-3 hBRAG cDNA, the T1 transcription mix (MBI Fermentas Inc., Flamborough, ON), and either T3 or T7 RNA polymerase (Promega) for 1 h at 37°C. The DNA template was removed with 2 units of DNase I for 15 min at 37°C. For subsequent translation reactions, an amino acid mixture without methionine, and [35S] methionine (Dupont) at 1 μCi/μl was added to the Flexi reticulocyte lysate translation mix (Promega, Madison, WI), all on ice. After incubation at 30°C for 1 h, tRNA was digested with RNase A. Translation reactions were divided into four equal aliquots, each with a final total volume of 20 μl. In particular, translations were performed either alone, in the presence of 1 μl (2 units) of canine pancreatic microsomal membranes (MBI Fermentas), with 1 μl microsomes and 40 μM of the N-glycosylation competitive inhibitor peptide NYT (Ac-Asn-Tyr-Thr-NH₂), or with 1 μl microsomes and digestion with 50 units of peptideN-glycosidase F (PNGaseF) for 2 h at 37°C subsequent to translation. Membrane integration of the translation products was determined by alkaline extraction of samples (10 μl) with 100 μl of ice-cold 0.1 M Na₂CO₃, pH 11.5, followed by recovery of the stripped microsomes by centrifugation (16, 000 g for 15 min.), direct resuspension in 2X electrophoresis sample buffer (New England BioLabs) supplemented with 1% 2-ME and 0.042 mM DTT, SDS-PAGE analysis, and autoradiography. Alternatively, prior to SDS-PAGE analysis, immunoprecipitations of the above translation reactions were performed with either pre-immune sera or N-terminal affinity-purified α-hBRAG antibodies under the same conditions as described above for steady-state and pulse labeling experiments. For cleavage of N-linked oligosaccharide groups in cellular proteins, lysates were solubilized in 1X electrophoresis sample buffer (New England BioLabs), diluted 1/5 in 0.5M
NaH₂PO₄ (pH 7.5) buffer supplemented with 1% NP-40, and treated with 100 units of PNGaseF at 37°C for 2 h.

Biotinylation of antibodies and intracellular/surface flow cytometry

α-BRAG 2422 affinity-purified, terminal bleed antibodies were biotinylated using D-biotin (Molecular Probes) using the method described previously [408]. For surface staining of BRAG expression on the K562 transfectants aliquots of 10⁶ cells were incubated on ice for 30’ with 10ug α-BRAG-biotin Ab. Cells were washed twice with cold PBS and incubated on ice for 30’ with 2ug/Streptavidin-FITC (Molecular Probes, Eugene, OR). Cells were washed twice with 4ml cold PBS and analysed using a FACScalibur flow cytometer (Becton Dickinson, San Jose, CA).

For intracellular staining of BRAG cells were incubated for 30’ on ice with 1ug a-BRAG-biotin Ab in 1% BSA/0.3% Saponin (Sigma Aldrich Canada Ltd, ONT, Canada)/PBS. Cells were washed twice in 2ml 0.1% Saponin/PBS and resuspended with 1ug Streptavidin-FITC in 1% BSA/0.3% Saponin/PBS. Cells were washed twice in 2 ml 0.1% Saponin/PBS and resuspended in 1% BSA/PBS and analysed immediately on the FACScalibur.

Cell surface biotinylation assay

Cells were washed with ice cold borate buffer (10 mM boric acid, 154 mM NaCl, 7.2 mM KCl, 1.8 mM CaCl₂, pH 9.0). EZ-Link NHS SS-Biotin (Pierce) at 0.8 mM in borate buffer for 15 minutes at 0°C was used to biotinylate the surface. The cells were then rinsed in 0.192 M glycine, 25 mM Tris, pH 8.3 solution to quench any unreacted reagent. The cells were then lysed with radioimmunoprecipitation assay (RIPA) buffer (1% deoxycholate, 1% Triton-X-100, 0.1% SDS, 0.15M NaCl, 1 mM EDTA, and 10 mM Tris-HCl, pH 7.5). An aliquot of the lysate was saved for Western blotting. ImmunoPure (Pierce) immobilized streptavidin (100 μl) was added to the lysate for 1 hour at 0°C to bind the biotinylated proteins. The supernatant was removed and an aliquot was saved for Western blotting. The streptavidin beads were washed three times with RIPA buffer. 1X Sample buffer (containing 5% 2-β-mercaptoethanol) was added to the beads and the
samples were boiled for 30 seconds to cleave the disulfide bond in the biotinylating reagent and release the captured proteins.

**Crosslinking and co-immunoprecipitation assays**

2x10^7 A8-6P cell aliquots/time point were resuspended at 5x10^6/ml of serum-free RPMI medium in 6-well plates and either left unstimulated, or cross-linked with 30 μg/ml of a polyclonal affinity-purified goat anti-human μ F(ab)’2 Ab (Tago Inc., Burlingame, CA) at 37°C for either 30 sec, 1 min, 3 min or 5 min. Cells were then washed in 9 ml of cold 1X PBS prior to direct solubilization in 1 ml of 1X sample loading buffer (New England BioLabs) for subsequent SDS-PAGE and immunoblot analysis of whole cell lysate fractions. Alternatively, 8x10^7 A8-6P cells unstimulated or cross-linked with 30 μg/ml anti-human μ F(ab)’2 for 3 min were prepared for immunoprecipitation with α-hBRAG, α-Hck, or α-lyn antibodies prior to SDS-PAGE and immunoblot analysis. Finally, for co-ligation studies, 2x10^7 A8-6P cells were ligated with various combinations of α-hBRAG, α-human μ, or α-hBRAG+α-human μ at various concentrations and for various durations (see figure legend for details). For blocking of hBRAG-specific crosslinking, hBRAG antibodies were pre-incubated with 10μg/ml competitor peptide UT952 overnight at room temperature. Cells were then solubilized in 4 ml of ice cold mild (co-immunoprecipitation) lysis buffer comprised of 150 mM NaCl, 50 mM Tris-HCl, pH 7.5, 5 mM EDTA, 5 mM EGTA, 5 mM NaFl, 1mM PMSF, 1mM sodium orthovenadate, and 1% Nonidet P-40, supplemented with the contents of a complete protease inhibitor mixture tablet (Boehringer Mannheim Corp., Indianapolis, IN). Immunoprecipitations were carried out as described above for pulse and steady-state labeling experiments. For blocking of antibody F_c receptors, an F_c blocking antibody directed against CD16/CD32 (FcγIII/II receptor block; kindly obtained from Kathy Siminovich, Samuel Lunenfeld Research Institute, Toronto, ON) was preincubated with A8-6P cells at 10 μg/ml for 1 h in a 37°C incubator prior to cell stimulation with α-hBRAG or IgG+α-human μ.
Chapter 4

SDS-PAGE

Whole cell lysates were sedimented by centrifugation at 14,000 g at 4°C for 10 min to remove cellular debris, sheared with a 21.5 gauge needle to eliminate chromosomal DNA, and boiled to denature proteins. Whole cell lysates, immunoprecipitates, or in vitro-synthesized proteins were electrophoretically resolved on 10% discontinuous SDS-PAGE Tris-Glycine pre-cast minigels (Novex, San Diego, CA) at 100 V for 2 h in Tris-Glycine running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS) using XCell II™ Mini Cell gel runner (Novex). Protein loads varied in various experiments, but in all cases either cell equivalents or 1 mg proteins were used and gels were stained with Comassie blue. Cell lines and tissue samples were washed in 1X PBS and solubilized in either non-denaturing buffer (2X New England BioLabs electrophoresis sample buffer alone, for runs in denaturing, non-reducing conditions, or in reducing buffer (2x NEB electrophoresis sample buffer supplemented with 1% 2-ME and 0.042 mM DTT) for runs in denaturing, reducing conditions. In each case, the Benchmark molecular pre-stained standard protein ladder (Gibco BRL, Grand Island, NY) was run in parallel. Following SDS-PAGE, radiolabeled proteins (from steady state labeling, pulse labeling, and in vitro translation experiments) were directly subjected to autoradiography for 24-48 h periods.

Immunoblotting and primary/secondary antibodies used

Unlabeled proteins were electrophoretically transferred to polyvinylidifluoride (PVDF) membranes (Millipore, Mississauga, ON) using the Xcell II™ blot module (Novex). The membranes were blocked by incubation in 0.25% gelatin, 10% ethanolamine, and 0.1 M Tris-HCl (pH 9.0) for 2 h at room temperature. The blocked nitrocellulose strips were then incubated with various dilutions of primary antibodies for 2 h at room temperature. The primary antibodies used in immunoblotting included: 4G10, a polyclonal anti-phosphotyrosine antibody (α-PTyr) kindly provided by Dr. Brian Druker (Oregon Health Sciences University, Portland, Oregon), monoclonal α-lyn and α-fyn antibodies kindly provided by Dr. Kathy Siminovich (Samuel Lunenfeld Research Institute, Toronto, Ontario), SC-72, a polyclonal α-human Hck antibody (Santa Cruz Biotechnology, Inc., Santa Cruz, CA), 6C8, a monoclonal hamster α-human Bcl2 antibody (Pharmingen), 386.12, a
monoclonal mouse α-human CD19 antibody (Pharmingen), mouse α-actin (Amersham, Arlington Heights, IL), pre-immune sera from the above described bleeds, crude antisera to hBRAG peptides, or affinity-purified α-hBRAG antibodies. Concentrations of antibodies used vary in different experiments and are detailed in the figure legends. In order to determine specific immunoreactivity of hBRAG antibodies i.e. detection of peptide blockable bands, duplicate blots in many experiments were run whereby primary antibodies were preadsorbed with their free relevant peptide (at various concentrations indicated in the figure legends) overnight at room temperature.

Bound primary antibodies were detected by then incubating membranes with either horseradish peroxidase (HRP) conjugated goat-α-rabbit or goat-α-mouse IgG (BioRad, Hercules, CA) at dilutions of 1:10000 and 1:5000, respectively, for 2 h at room temperature. For blotting of immunoprecipitates with antibodies against the human proteins Hck and hBRAG, HRP-conjugated protein A was used as a second step antibody in immunoblotting to decrease background caused by IgG cross-reactivity. The buffer used for washes and incubations was 0.25% gelatin, 0.05% Nonidet P-40, 0.015 M NaCl, 5mM EDTA, and 0.05 Tris-HCl (pH 7.5). After thorough washing, western blots were developed using the Renaissance enhanced chemiluminescence detection system (Dupont NEN) and blots were exposed for 5-30 s using BioMax™ MR autoradiography film (Eastman Kodak, Rochester, NY).
4.4 RESULTS AND DISCUSSION

4.4.1 Characterization of hBRAG antibody specificity

To begin biochemical characterization of the hBRAG protein, affinity-purified polyclonal antibodies were generated against N and C-terminal epitopes of the hBRAG protein (see Materials and Methods). In the experiments presented in this chapter, the binding specificity of these antibodies were tested in various contexts, including against BSA-conjugated peptides, whole cell lysates of endogenously-expressing hBRAG+ cell lines and stably-transfected lines, and in vitro synthesized products. However, prior to use of α-hBRAG antibodies in immunoblotting and immunoprecipitation assays, specific reactivity of the various bleeds of N and C-terminal hBRAG antisera and corresponding purified antibodies against the hBRAG peptides was tested in Western blotting on BSA-conjugate membrane strips prepared by electrophoretically fractionating 10 μg BSA-conjugated peptides by SDS-PAGE and transfer to PVDF membranes (see Appendix, Figure A.3). Based on these results, both N and C-terminal affinity purified antibodies showed strong reactivity towards their corresponding peptides relative to pre-immune controls, with the strongest reactivity seen in antibodies derived from terminal bleeds.

Peptide competition assays using whole cell lysates from the endogenous hBRAG+ line A8-6P were then used to test the binding of crude sera and corresponding affinity purified antibodies from initial bleeds on Western blots (Figure 4.1A). These assays showed that relative to pre-immune serum controls, N-terminal antibodies detect two specific i.e. peptide blockable antigens: a ~52 kDa major band and a minor ~63 kDa antigen. Based on the size difference of the minor product, it can be hypothesized that it may represent an inefficiently glycosylated isoform of the hBRAG protein. Using the N-terminal, affinity-purified terminal bleed α-hBRAG antibodies, the same ~52 kDa and ~65 kDa specific peptide blockable bands found in A8-6P were also observed in the stable expressing K562 8-3 pCEP4-1 hBRAG transfectant, but not in the mock-transfected k562 control (Figure 4.1B). The C-terminal antibodies, on the other hand, bind to multiple different proteins and/or multiple hBRAG protein isoforms, the results which for the sake of simplification, will not
Figure 4.1. Immunochemical characterization of rabbit polyclonal antiserum.  
A. Western blot analysis of rabbit polyclonal antiserum reactivity to membrane strips prepared from endogenous hBRAG+ A8-6P cell line lysates. Two specific immunoreactive products (a major, ~52 kDa product and a minor, ~63 kDa product; denoted by *) were detected by hBRAG 13th week bleed rabbit polyclonal antiserum (lot 2422) raised against the hBRAG N-terminal peptide antigen UT952, or by corresponding affinity purified α-hBRAG antibodies. Products were considered specifically immunoreactive if not found in pre-immune sera and blockable in peptide competition assays i.e. incubating multiple duplicates of membrane strips with each antibody after overnight pre-incubation with increasing concentrations of competing peptide antigen UT952. Membrane strips were incubated with antibody alone (denoted as "-"*) or with increasing amounts of UT952 (denoted by triangles). Lanes are as follows: lane 1; incubation with pre-immune serum at a 1:100 dilution, lanes 2, 3; incubation with pre-immune serum at a dilution of 1:100 and pre-adsorbed with 100 ng and 10 μg of UT952, respectively, lane 4; incubation with crude terminal bleed hBRAG rabbit polyclonal antiserum at a 1:100 dilution, lanes 5-8; incubation with crude hBRAG antiserum at a 1:100 dilution, lane 9; incubation of 1 μg/ml affinity purified α-hBRAG antibodies alone, and lanes 10-13, incubation of 1 μg/ml affinity purified α-hBRAG antibodies pre-adsorbed with 10ng, 100 ng, 1 μg, 10 μg/ml UT952, respectively. 
B. Western blot analysis of membrane strips prepared from either untransfected K562 lysates (lanes 1 and 2) or from stable hBRAG transfectant K562 8-3 pCEP4-4 lysates (lanes 3 and 4). Detection of the specific immunoreactive ~52 and 63 kDa bands (denoted by *) was by incubation with 1 μg/ml affinity purified α-hBRAG antibodies either alone (Lanes 1 and 3) or in the presence of pre-adsorbed 10 μg/ml of UT952 (lanes 2 and 4). Note the greater relative intensity of the specific 65 kDa product relative to the 52 kDa product in K562 transfectant compared to that seen in endogenous lysates (Figure 1A and B; Figure 3; see text for details). In both panels A and B, membrane strips all contain equal protein loads (40 μg) and lysates have been run under denaturing, non-reducing SDS-PAGE conditions.
A

[Image of protein gel electrophoresis showing bands at various molecular weights (kDa) for pre-immune, crude antisera, and affinity purified samples.]

- 87
- 64
- 52
- 39
- 26
- 21

1 2 3 4 5 6 7 8 9 10 11 12 13

B

[Image showing protein bands for K562 and K562 8-3 samples with and without blocking peptide.]

- 194
- 120
- 87
- 64
- 52
- 39
- 26

1 2 3 4
be discussed further here but rather will be addressed in the general discussion (also see Appendix, Figures A.3 and A.4 for specific details)

4.4.2 Identification of hBRAG as a glycosylated, integral membrane protein

To test whether the larger specific band seen in Western blots represent the glycosylated isoform of the hBRAG product, two approaches were taken. First, a cell-free system consisting of a reticulocyte lysate supplemented with microsomal membranes was used to assay the N-glycosylation status of the hBRAG protein in vitro (Figure 4.2A). Consistent with structural predictions of the hBRAG protein (without post-translational modifications), I had demonstrated in Chapter 3 that the in vitro translated hBRAG protein is present as a doublet of ~50 and 55 kDa bands respectively, the smaller band possibly due to an alternative downstream translational initiation site in the hBRAG sequence N-terminus (Chapter 3, Figure 3.4C). The same doublet in cell free translations was also seen in the current cell-free translations alone (Figure 4.2A, lane 2). However, N-glycosylation at a single site will produce a uniform increase of about 2.5 kDa in the molecular weight of the protein corresponding to the cotranslational attachment of the high mannose oligosaccharide, which only occurs in the microsomal lumen [339]. Consistent with the structural prediction of 4 potential N-glycosylation sites in the hBRAG extracellular domain, the in vitro synthesized hBRAG protein was found to shift to a ~10 kDa larger species (Figure 4.2A, lane 3). To further extend the above results, α-hBRAG antibodies were used to specifically immunoprecipitate in vitro synthesized hBRAG proteins, either alone or in the presence of microsomes. It was shown that the N-terminal hBRAG affinity purified antibody detects the full-length hBRAG protein doublet expressed after cell free in vitro translation (Figure 4.2A, lane 6). Additionally, this antibody also recognizes the post-translationally modified in vitro translated proteins from membrane fractions (Figure 4.2A, lane 7).

Two independent experiments were performed in parallel to confirm that the shift in mobility was specifically due to N-glycosylation rather than to other possible post-translational modifications. I used either the N-glycosylation-cleavage specific enzyme N-glycosidase F, and in addition used a competitive peptide inhibitor of N-glycosylation, the acceptor peptide Ac-Asn-Tyr-Thr-NH₂
### A

<table>
<thead>
<tr>
<th></th>
<th>whole</th>
<th>IPS: α-hBRAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>S</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>S</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>S</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>S</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>8-3 cDNA</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>microsomes</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PNGase F</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>NYT</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>kDa</th>
<th>194</th>
<th>120</th>
<th>87</th>
<th>64</th>
<th>52</th>
<th>39</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### B

<table>
<thead>
<tr>
<th></th>
<th>A8-6P</th>
<th>K562</th>
<th>K562 8-3 pCEP4-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>8-3 cDNA</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>microsomes</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>PNGase F</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>kDa</th>
<th>120</th>
<th>87</th>
<th>64</th>
<th>52</th>
<th>39</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.2. N-glycosylation of in vitro-translated hBRAG cDNA, and of endogenous and transfected hBRAG products in cellular lysates. A. N-glycosylation of the in vitro-translated hBRAG cDNA and recognition of glycosylated and unglycosylated hBRAG isoforms by α-hBRAG antibodies. Autoradiogram of the [35S] methionine-labeled cell-free translated hBRAG gene products either directly resolved by SDS-PAGE analysis (lanes 1-5), or immunoprecipitated with affinity purified α-hBRAG antibodies from terminal bleed antiserum raised against the N-terminal hBRAG peptide prior to SDS-PAGE UT952 (2422; lanes 6-9). The open arrow shows the unglycosylated hBRAG doublet product and the closed arrow shows the ~70 kDa glycosylated product shown by a closed arrow (see text for details of relative amounts of glycosylated and unglycosylated products in whole and in immunoprecipitated cell-free fractions). Numbers to the left indicate the migration positions of the protein standards. Lane 1 represents a negative control in which no cDNA was added. Translation reactions were carried out either without microsomes (lanes 2 and 6), with microsomes only (lanes 3 and 7), with microsomes and the competitive inhibitor NYT (lanes 4 and 8), or with microsomes only and digestion with PNGaseF after translation (lanes 5 and 9). The designation “S” indicates the linearized pBluescript 8-3 hBRAG that was transcribed in vitro, using the T3 promoter, to generate a sense RNA, whereas “AS” indicates in vitro transcription of 8-3 hBRAG with the T7 promoter to generate an anti-sense RNA. B. PNGaseF treatment of endogenous hBRAG-expressing A8-6P cellular lysates or transfected/untransfected non-endogenous hBRAG-expressing K562 cellular lysates. Lysates were resolved on SDS-PAGE gels under reducing conditions as detailed for Figure 1, and immunoblotted using affinity purified N-terminal α-hBRAG antibody 2422 at 1:1000 dilutions. The (+) signs (lanes 2, 4, 6) represent samples treated with 5000 U PNGaseF for 1h at 37°C; (-) signs (lanes 1, 3, 5) denote untreated samples. The putative glycosylated and unglycosylated isoforms (~65 kDa and 52 kDa, respectively) are shown as open and closed arrows, respectively. Note that consistent with the structural prediction of four N-glycosylation sites, the difference in sizes between the glycosylated products in lysates and in cell-free translations is ~10 kDa.
(NYT). With either addition of NYT or treatment with N-glycosidase F, the N-glycosylation-mediated shift seen in the presence of microsomes was not observed (Figure 4.2A, lanes 3, 4). Furthermore, in vitro translated products immunoprecipitated with α-hBRAG antibodies also demonstrated a specific loss of the glycosylated hBRAG isoform from the microsomal fraction, either with NYT pre-incubation or with N-glycosidase F post-treatment (Figure 4.2A, lanes 8, 9). Taken together, the results suggest that hBRAG is a membrane-integrated glycoprotein, as has been predicted based upon the aa sequence.

Based on densitometric analysis, the upper (~60 kDa) band of the unglycosylated hBRAG cell-free translated product doublet is 2.1 and 2.4 fold more intense than the bottom (~55 kDa) band, prior to the addition of microsomes in whole and immunoprecipitated fractions (Figure 4.2A, lanes 2 and 6, respectively). Upon microsome addition, however, the relative amounts of the upper and lower bands are roughly equivalent (Figure 4.2A, lanes 3 and 7), but the upper band returns to increased amounts relative to the lower band either upon PNGaseF or NYT treatment (Figure 4.2A, lanes 4, 5, 8, 9), suggesting the larger species is selectively glycosylated. It is unlikely that this selective glycosylation is due to alternative translational initiation codon usage, since the potential alternative translation initiation site with a Kozak-like sequence is about ~40 aa C-terminal to the first site, and therefore also encodes a protein with four N-glycosylation sites in its extracellular region. One intriguing possibility, however, is that hBRAG, at least in vitro, may be capable of inserting itself in the membrane in two orientations, either as a type II, or in the inverted orientation, as a type I, and as such perhaps only in one orientation is it glycosylated. The insertion event itself may depend on the cytoplasmic region and its relative length such that a shorter cytoplasmic tail may allow improper insertion of the protein as a type I, rather than type II protein, and the extracellular domain with N-glycosylation sites now is the cytoplasmic domain and will not be accessible to N-glycosylation. In this context, alternative initiation site usage may therefore indirectly determine glycosylation status.

A second test for hBRAG glycosylated and unglycosylated isoforms was to directly assess the glycosylation status of either the endogenous hBRAG or transfected cellular hBRAG protein. This
was accomplished by treatment of lysates with PNGaseF prior to immunoblot analysis with N-terminal specific α-hBRAG antibodies (Figure 4.2B). After PNGaseF treatment, the minor product is selectively lost, and relative to the untreated sample, there is more 52 kDa product in both the endogenously-expressing cell line A8-6 P and in the hBRAG transfectant K562 8-3 pCEP4-4 (Figure 4.2B, lanes 2 and 6). The low level of glycosylated hBRAG expression seen in these lanes, in Figures 4.1B and D (seen only with longer exposures) and with affinity-purified, but not crude antisera, is consistent with the relatively low percentage of glycosylated product (~25% as assessed by densitometric analysis of glycosylated/unglycosylated product ratios) seen in immunoprecipitates of cell-free translations (Figure 4.2, lane 7). This could be due to two possibilities: either the actual glycosylation of this protein is not efficient and there may exist a large pool of unglycosylated hBRAG, or alternatively, the N-terminal α-hBRAG antibody is inefficient at recognizing a possibly different glycosylated conformation. The latter possibility is consistent with the similar intensity of glycosylated product in α-hBRAG immunoprecipitations and whole cell lysates (Figure 4.2A, compare lanes 3 with 7). Interestingly, relative to peptide competitions with the endogenously-expressing hBRAG cell line A8-6P, a higher glycosylated:unglycosylated hBRAG ratio is seen in stable hBRAG transfectant K562 8-3 pCEP4-4 lysates, indicating either the transfected hBRAG is more efficiently recognized by antibody or is more efficiently synthesized by cellular machinery. It is also formally possible that the weak glycosylation form seen in endogenous lysates could be the real antigen, whereas the more ubiquitous 52 kDa antigen is not the unglycosylated antigen, but a cross-reacting, higher abundance protein. However, this latter possibility is not consistent with the in vitro translation data, and the shift in size seen after PNGaseF treatment of A8-6P and K562 8-3 pCEP4-4 lysates. Nevertheless, the fact that the lower molecular weight hBRAG isoform appears to be differently post-translationally modified could potentially mean it serves a different function and/or does not have a motif that is present in the longer protein which alters its cellular distribution i.e. intracellular versus cell surface expressed.
4.4.3 Analysis of hBRAG protein expression in lymphoid cell lines and tissues and identification of hBRAG disulfide-linked multimers

Almost all type II transmembrane proteins expressed in lymphoid cells and/or with immunologic function (many with important signaling functions in T-B interactions) exist on cell surfaces not only as glycosylated receptors, but as higher order structures, either as homodimers (CD69, CD72, LY49, CD94), heterodimers (CD94/NKG2), or homotrimers (CD23; reviewed in [351]).

To test if hBRAG, like other type II proteins, could form higher order structures, lysates from various cell lines and tissues were prepared and run under denaturing reducing and non-reducing conditions prior to immunoblotting with α-hBRAG antibodies, either in the presence or absence of blocking peptide (in order to assess the specificity of products resolved under both conditions). As expected, under reducing conditions, the specific 52 kDa major antigen was expressed in B cell lines, but not in the non-lymphoid line HeLa or the myeloid cell line U937 (Figure 4.3A). Under the particular conditions of this assay and the relatively high background and short exposure times, the glycosylated product was not detected. Nevertheless, under non-reducing conditions, in addition to resolving the single polypeptide hBRAG antigen, a larger disulfide-linked specific molecule of ~120 kDa could also be resolved in B cells. Based on its molecular weight, this molecule is not likely to represent an unglycosylated hBRAG homodimer, but is probably either a glycosylated hBRAG homodimer or an unglycosylated/glycosylated heterodimer. A predominant ~80 kDa band was also seen in the pre-B cell line 697 under reducing conditions, but is likely not a hBRAG heterodimer, since it is not competed out by peptide, and more likely represents a 697-specific cross-reacting antigen. In tissues, a lot more cross-reactivity to non-specific i.e. non-peptide blockable bands, was seen, but both the highly-expressed 52 kDa and the lower-expressed 63 kDa specific single polypeptide products seen previously under denaturing conditions in immunoblots of B cell line lysates were seen in bone marrow and PBL, but not thymus, heart or placenta unfractionated tissues. Under non-denaturing conditions, the same specific products were seen in the same tissues, but additional, larger dimer and possibly trimer disulfide-linked molecules of ~120, 150, and 180 kDa (representing either hetero/homo glycosylated hBRAG dimers and/or unglycosylated/glycosylated hBRAG hetero/homotrimers) were also seen.
Figure 4.3. Western blot analysis of various cell lines and tissues with α-hBRAG antibodies reveals hBRAG dimer formation.

A. Western blot analysis of protein lysates prepared from 2x10^7 cell equivalents of various cell lines fractionated by SDS-PAGE under either non-reducing conditions (lanes 1-8) or reducing conditions (lanes 9-18). Cells include the human pro-B, pre-B and mature B cell lines REH, 697, and A9-6P, respectively, the human fibroblast cell line HeLa, and myelocytic human cell line U937. B. Western blot analysis of 1 mg protein lysates from various unfractionated human tissues resolved by SDS-PAGE under either non-reducing conditions (lanes 1-10) or reducing conditions (lanes 11-20). In both panels A and B, after proteins were fractionated by SDS-PAGE, they were transferred to a PVDF membrane, and affinity-purified antibodies from terminal bleeds of crude hBRAG antisera raised against N-terminal peptide UT952 at a concentration of 1 μg/ml either alone (-) or in the presence of excess (500 μg/ml) competing peptide UT952 (+) were used to stain the blots, followed by detection with a chemiluminescence detection system. The positions of molecular weight standards (expressed in kilodaltons) are indicated on the left. The positions of the specifically immunoreactive bands that are competed out by N-terminal peptide are indicated with arrows.
To extend the above immunoblotting results of gels run under non-denaturing conditions, and to compare hBRAG isoforms during biosynthesis (prior to, and after post-translational modification), endogenous hBRAG-expressing A8-6P B cells were either metabolically pulse or steady-state labeled, immunoprecipitated with α-hBRAG antibodies, and run under non-reducing conditions (Figure 4.4). As in reduced gel immunobLOTS, larger bands likely representing multiple hBRAG higher order isoforms, (possible homo or heterodimers) were seen in both steady and pulse-labeled A8-6P immunoprecipitates (Figure 4.4). Although the specific product sizes are similar between steady state-labeled immunoprecipitates (Figure 4.4B) and whole cell lysate products in western blot assays (Figure 3), they are slightly larger in immunoprecipitations of pulse-labeled cellular proteins (Figure 4.4A). This variation cannot be due to alternative splicing because only one transcript is seen in all cell lines assessed (Chapter 3, Figure 3.5). However, one possibility is that hBRAG may be a highly unstable protein which is particularly susceptible to proteolysis, such that the antisera are detecting various hBRAG degradation isoforms, depending on the conditions of the assay. Overall, the results of these assays are consistent with the structural prediction of hBRAG being expressed in B cells at least partly as a disulfide-linked dimer, similar to other lymphocyte-specific cell-surface expressed type II transmembrane receptors.

4.4.4. hBRAG intracellular and surface isoforms

To assess if hBRAG exists as a cell surface protein or an intracellular protein, flow cytometry of K562 cells stably transfected with hBRAG (8-3) cDNA, was performed using primary biotinylated versions of the hBRAG N-terminal antibodies used in Western blotting and immunoprecipitation assays (Figure 4.5A). To control for specific staining, the same assays were performed in the presence of blocking peptide. Relative to the K562 mock transfectant, positive surface and intracellular staining was seen in the K562 hBRAG transfectant. Furthermore, this appeared to be specific for hBRAG, as the same shift was not seen in the presence of peptide blocking or when using an irrelevant biotinylated primary IgG (Figure 4.5A).
Figure 4.4. Analysis of hBRAG biosynthesis as assessed by immunoprecipitations of metabolic steady and pulse-labeled A8-6P whole cell lysates.

A. Autoradiogram of SDS-PAGE fractionated lysates from 2x 10^6 A8-6P cells that were pulse-labeled with [35S] methionine, and immunoprecipitated with either pre-immune sera at 1:200 and 1:40 dilutions (lanes 1 and 2, respectively), 1:1000, 1:200, 1:40 and 1:10 dilutions of crude N-terminal hBRAG antisera (lanes 3-6, respectively), or with a 1:10 dilution of crude N-terminal hBRAG antisera pre-incubated with excess (500 μg/ml) blocking peptide UT952 (lane 7). B. Autoradiogram of 1 x 10^6 A8-6P cells that were steady state-labeled with [35S] methionine, and immunoprecipitated with either pre-immune sera at a 1:10 dilution (lane 2), crude N-terminal hBRAG antisera at a 1:10 dilution without or with excess (500 μg/ml) blocking peptide UT952 (lanes 3 and 4, respectively), or with affinity purified α-hBRAG antibodies at a 1:100 dilution with or without excess (500 μg/ml) blocking peptide UT952 (lanes 5 and 6, respectively). Lane 1 represents the precleared labeled lysate fraction. For both panels A and B, the positions of molecular weight standards (in kilodaltons) are indicated on the left, and arrows to the right identify the position of specifically immunoreactive bands i.e. blockable by immunizing antigen and not found in pre-immune sera.
Figure 4.5. Analysis of intracellular and cell surface–expressed hBRAG isoforms.

A. Flow cytometry of K562 cells stably transfected with empty vector or with vector containing hBRAG full-length (8-3) cDNA, using biotinylated affinity purified α-hBRAG antibodies (2422) from terminal bleed antiserum raised against the N-terminal hBRAG peptide UT952. As negative and positive controls, antibodies were pre-incubated with specific blocking or irrelevant peptide controls, respectively. B. Western blot analysis of intracellular and cell-surface hBRAG expression. Intracellular protein fraction lysates (designated “I”) or biotinylated, cell-surface protein supernatants (designated “S”) from A8-6P K562 empty vector controls (lanes 5 and 6, respectively), or K562 hBRAG transfectants (lanes 7 and 8, respectively) were assessed by immunoblotting using the 2422 affinity purified α-hBRAG antibodies 2422 at a concentration of 1μg/ml (see Materials and Methods for details of cell surface biotinylation assay). The same assay was used to assess cell surface supernatants or intracellular lysates of unfractionated tonsils (lanes 9 and 10, respectively) or purified tonsillar B cells (lanes 11 and 12, respectively). As controls for intracellular and plasma membrane fractionation, bcl2 (lanes 1, 2) and CD19 protein expression (lanes 3, 4) from A8-6P fractions were assessed under the same assay conditions. The open and closed arrow show the putative glycosylated and unglycosylated hBRAG products, respectively. The positions of molecular weight standards (expressed in kilodaltons) are indicated in the middle.
A

- α-hBRAG-FITC alone
- α-hBRAG-FITC + irrelevant peptide
- filled α-hBRAG-FITC + UT952 peptide

intracellular

K562 pCEP4

surface

K562 8-3 pCEP4-4

B

<table>
<thead>
<tr>
<th>Immunoblot</th>
<th>Controls</th>
<th>α-CD19</th>
<th>α-bcl2</th>
<th>α-hBRAG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subcellular fraction</td>
<td>I S I S</td>
<td>I S I S</td>
<td>I S I S</td>
<td></td>
</tr>
<tr>
<td>A8-6P K562 pCEP4 K562 8-3 pCEP4-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole tonsil Purified B cells</td>
<td>I S I S</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- 1 2 3 4 5 6 7 8 9 10 11 12
To extend the results seen by flow cytometry in transfectants, I performed cell surface biotinylation assays in both the above transfectants and in purified human tonsillar B cells (see Materials and Methods; Figure 4.5B). Consistent with the flow cytometry data, both a cell-surface expressed and intracellular hBRAG fraction is seen in hBRAG transfectants, whole tonsillar fractions, and in purified B cells, but not in mock transfectants. In the intracellular fraction, more unglycosylated product relative to glycosylated product is present in endogenously-expressing cells, whereas a relatively equivalent level of both isoforms in this pool is seen in the K562 hBRAG transfectants, results which are similar to that seen in cell-free translations and immunoblots of whole cellular lysates (Figures 4.1-4.3). Interestingly, both the putative unglycosylated and glycosylated products (~52 and 63 kDa, respectively) are seen in the intracellular fraction, but only the larger, glycosylated product is seen in the surface fractions (Figure 4.5B). Furthermore, in all lysates assayed using this particular assay, a large percentage of the post-translationally modified hBRAG protein is retained intracellularly relative to that expressed on the plasma membrane, a finding that is consistent with the structural prediction of a short hBRAG transmembrane region (Chapter 3, Figure 3.3). Finally, these results show that there is more surface and intracellular product in enriched B cells than in whole tonsil, a result that is consistent with preferential expression of this product in human B cell lines and tissues (Chapter 3; Figures 3.5 and 3.6). Overall, these results suggest that hBRAG is expressed as both a low, but detectable glycosylated population on the cell surface, and a larger, intracellular fraction of glycosylated and unglycosylated isoforms.

With the exception of invariant chain, all the above-mentioned type II TM protein multimers with characterized function in the immune system are preferentially expressed on the cell-surface. Invariant chain, which is a predominantly intracellularly-expressed protein, without its ER retention signal, can also be expressed on the surface at low, but detectable levels. The low, but detectable levels of hBRAG in the above cell surface staining and biotinylation assays (Chapter 4, Figure 4.5) is therefore consistent with the structural prediction that hBRAG, which lacks such a signal, should at least partly leak out to the plasma membrane. Consistent with cell-surface expressed hBRAG is the fact there are candidate signaling motifs in its 81 aa cytoplasmic N-terminus including consensus
cAMP and PKC phosphorylation sites and palmitylation residues (See Chapter 3, Figure 3.4), shared by various other signaling receptors and intermediates, such as G protein-linked receptors, G proteins, and Src family kinases including Src, Blk, Lyn, Yes, Hck, Fyn, and Lck (reviewed in [410]).

4.4.5 hBRAG is associated with phosphorylated proteins in unstimulated cells that are dissociated and/or dephosphorylated in response to BCR signaling

To begin assessing if hBRAG may be part of a signaling cascade, (either as a signaling receptor or intermediate), proteins were immunoprecipitated with hBRAG antibodies under mild conditions and immunoblotted with α-phosphotyrosine antibodies. As can be seen in Figure 4.6A, hBRAG was found to associate with phosphorylated proteins in co-immunoprecipitations of unstimulated A8-6P cells. I then assessed to examine the possibility that hBRAG may also be functionally associated with BCR signaling, I assessed changes in physical interactions with other phosphorylated proteins upon BCR ligation. To do this, proteins from OCI LY8-C3P cells that had been cross-linked with saturating concentrations of polyclonal anti-μ antibody were immunoprecipitated under mild conditions, and subsequently immunoblotted with anti-phosphotyrosine antibodies. From previous experiments, I have observed optimal tyrosine phosphorylation in whole cell lysates using 10 μg/ml F(ab)’2 anti-μ for 3 minutes [75, 264, 411]. Surprisingly, α-phosphotyrosine immunoblots of immunoprecipitates revealed that a ~55-59 kDa protein doublet associated with hBRAG in resting cells may in fact either dissociate from hBRAG or alternatively, be selectively dephosphorylated in response to BCR ligation (Figure 4.6A). A weaker set of larger products also seems to be associated with hBRAG which are also either dephosphorylated or disassociated upon BCR ligation, and could possibly represent multimers of the major ~55-59 kDa doublet products (Figure 4.6A). The specificity of this doublet (and the larger, minor products) detected in unstimulated B cells under co-immunoprecipitating conditions appears to be sufficient to effectively eliminate the phosphorylated/dephosphorylated hBRAG-associated antigen with competing N-terminal peptides (Figure 4.6A). It is unlikely that that this is hBRAG itself, based on the molecular weight of the specific hBRAG band that is detected when the same blot is stripped and reprobed with the N-terminal hBRAG antibody (Figure 4.6B).
Figure 4.6. Detection of potential hBRAG-associated phosphorylated proteins in resting, but not BCR-ligated A8-6P lysates immunoprecipitated with α-hBRAG antibodies.

A. Anti-phosphotyrosine (α-P Tyr) Western immunoblots showing the effect of BCR ligation on protein tyrosine phosphorylation in whole cell lysates or α-hBRAG immunoprecipitated lysates prepared from hBRAG* mature B cell line variant A8-6P. Whole cell lysates alone from 2 x 10⁷ unstimulated A8-6P cells (denoted as “0”) or from 2 x 10⁷ A8-6P cells cross-linked with 30 μg/ml F(ab)² anti-μ for either 30 sec, 1 min, 3 min, or 5 min were subjected to SDS-PAGE, transferred to PVDF membranes, and probed with a polyclonal anti-phosphotyrosine antibody. Alternatively, lysates prepared from 2 x 10⁷ unstimulated A8-6P cells or A8-6P cells cross-linked with 20 μg/ml F(ab)² anti-μ for 3 min were immunoprecipitated under mild (co-immunoprecipitating) conditions (i.e. 1% TritonX-100, 150 mM NaCl, 50 mM Tris-HCl, pH 8, 5 mM EDTA, 5 mM EGTA, 5 mM NaF, 1 mM sodium orthovanadate, 1 mM BSA, and protease inhibitors) with either irrelevant affinity purified rabbit IgG at 10 μg/ml (denoted rIgG), or N-terminal 2422 α-hBRAG antibodies affinity purified from terminal bleeds at 10 μg/ml dilutions (denoted α-hBRAG) were subjected to SDS-PAGE, transferred to PVDF membranes, and probed with 1.5 μg/ml polyclonal 4G10 anti-phosphotyrosine antibody. The position of the putative major ~53/56 kDa hBRAG co-immunoprecipitating doublet species that is dephosphorylated upon BCR stimulation is denoted by an arrow, and the positions of molecular weight standards are indicated in the middle. B. The same Western blots stripped and re-immunoblotted α-hBRAG antibodies. The same affinity purified 2422 α-hBRAG antibodies used in immunoprecipitations were used, but at a concentration of 1 μg/ml. For both A and B, the blots were also probed without lysate, denoted as “(-)”, to control for cross-reactivity of the mouse α-rabbit second step HRP-conjugated reagent. 200 μg/ml of irrelevant rabbit IgG alone added to immunoprecipitation buffer was also used as a negative control. To test for specificity of bands under the described co-immunoprecipitating conditions, 10 μg/ml competing N-terminal peptide U-T-952 was used (denoted as “+UT952”); for comparison, an unstimulated sample without peptide blocking was run adjacenty (denoted as “–UT952”). The position of the putative major ~52 kDa hBRAG unglycosylated species recognized by affinity purified 2422 is denoted by an arrow. Positions of molecular weight standards are indicated at the left.
A

<table>
<thead>
<tr>
<th>BCR stimulation</th>
<th>Whole cell lysates</th>
<th>IP: rlgG</th>
<th>IP: α-hBRAG</th>
<th>pre-cleared fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>30&quot;</td>
<td>1'</td>
<td>3'</td>
<td>5'</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>3'</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

kDa

- 194
- 120
- 87
- 64
- 52
- 39
- 26

B

blot: α-P Tyr

- 52

blot: α-hBRAG
Various candidate doublet molecules that are constitutively phosphorylated in B cells in the ~55-59 kDa molecular weight range include three B-cell protein tyrosine kinases (PTKs) known to be associated with BCR signaling: Fyn, Lyn, and Hck, the latter two which exist as doublet isoforms of 56/59 and 53/56 kDa, respectively. To assess the possibility that hBRAG associates with any or a combination of these candidate proteins, antibodies against Lyn, Fyn, and Hck were used in co-immunoprecipitation experiments of the same A8-6P lysates under the same conditions as done for the original hBRAG co-immunoprecipitations shown in Figure 4.6. The parallel experiment, co-immunoprecipitation with hBRAG was also performed, and immunoprecipitates were resolved on SDS-PAGE and immunoblotted with either α-phosphotyrosine, α-hBRAG and either α-lyn, α-fyn, or α-Hck antibodies (Figure 4.7). No association with Lyn (Figure 4.7A) or Fyn (data not shown) was detected, but hBRAG demonstrated low level association with Hck (for Hemopoietic Cell Kinase) in both α-Hck immunoprecipitates immunoblotted with α-hBRAG and in α-hBRAG immunoprecipitates immunoblotted with α-Hck antibodies (Figure 4.7B). Based on the intensity of the associated proteins detected in α-phosphotyrosine immunoblots of α-hBRAG immunoprecipitates, Hck is likely not the only hBRAG associated antigen being dephosphorylated or dissociated upon BCR ligation in A8-6P B cells (Figure 4.7A). Also noteworthy is that the Hck fraction associated with hBRAG in resting A8-6P cells is likely de-phosphorylated rather than dissociated, since comparable levels of Hck are seen in α-Hck immunoblots of α-hBRAG immunoprecipitates either in the presence or absence of BCR ligation (Figure 4.7B). This assumes that Hck at least partially corresponds to the species that is selectively de-phosphorylated in BCR-stimulated hBRAG immunoprecipitates. Finally, it is noteworthy that hBRAG itself is not phosphorylated/dephosphorylated itself, at least in the fraction which associates with Hck in α-Hck immunoprecipitates. This, however, does not rule out that hBRAG is phosphorylated/dephosphorylated as part of the 60 kDa fraction in whole cell lysates. Further experiments will have to be performed to examine this possibility.

It is potentially interesting that hBRAG, as a possible surface receptor may associate with Hck, since this Src PTK family member expressed is predominantly in myeloid and B cell lineages [412, 413]. Although involvement of Hck in myeloid lineage signaling pathways has been fairly well
Figure 4.7. Co-immunoprecipitation of A8-6P lysates with anti-hBRAG antibodies, or antibodies against candidate hBRAG-associated, doublet B-cell PTKs Lyn and Hck.

Lysates were prepared from $2 \times 10^7$ A8-6P cells that were unstimulated A8-6P (denoted as "-") or cross-linked with 20 µg/ml F(ab)$_2$ anti-µ for 3 minutes (denoted as "+"). Immunoprecipitation conditions, and cell and antibody concentrations used are detailed in Figure 6 and in the Materials and Methods section. Immunoprecipitates were resolved in duplicate on SDS-PAGE gels under denaturing conditions, transferred to PVDF membranes, and probed with either α-hBRAG N-terminal affinity-purified antibodies or antibodies directed against either the Lyn or Hck PTKs.
### A

<table>
<thead>
<tr>
<th>Whole Cell Lysates</th>
<th>Immunoprecipitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCR ligation</td>
<td>α-rigG</td>
</tr>
<tr>
<td></td>
<td>α-hBRAG</td>
</tr>
<tr>
<td></td>
<td>α-lyn</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>kDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
</tr>
<tr>
<td>87</td>
</tr>
<tr>
<td>64</td>
</tr>
<tr>
<td>52</td>
</tr>
</tbody>
</table>

**α-pTyr immunoblot**

52 kDa

**α-hBRAG immunoblot**

53/56 kDa

**α-lyn immunoblot**

1 2 3 4 5 6 7 8

### B

<table>
<thead>
<tr>
<th>Whole Cell Lysates</th>
<th>Immunoprecipitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCR ligation</td>
<td>α-rigG</td>
</tr>
<tr>
<td></td>
<td>α-Hck</td>
</tr>
<tr>
<td></td>
<td>α-hBRAG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>kDa</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
</tr>
<tr>
<td>87</td>
</tr>
<tr>
<td>64</td>
</tr>
<tr>
<td>52</td>
</tr>
</tbody>
</table>

**α-pTyr immunoblot**

52 kDa

**α-hBRAG immunoblot**

56/59 kDa

**α-hck immunoblot (short exposure)**

56/59 kDa

**α-hck immunoblot (long exposure)**

1 2 3 4 5 6 7 8
characterized, its precise role in the BCR signaling pathway is not well understood. However, evidence for its involvement in this process comes from several observations. First, upon BCR ligation, Hck can be phosphorylated and activated, demonstrating *in vitro* kinase activity and phosphorylation of various *in vivo* BCR signaling-intermediate targets including p120 (Cbl), Bcr-abl, and RAS-GTPase-activating protein (GAP; reviewed in [19, 414]). Second, Hck has been shown to interact with other BCR or pre-BCR components/signaling intermediates including Ig-β and Ig-α (via its SH2 domains), and Bruton’s Tyrosine Kinase (BTK) and the Bcr-abl tyrosine kinase (primarily via its SH3 domains; ([415, 416]; reviewed in [417]).

The physical and functional interaction of Hck with c-abl, and possibly with hBRAG is intriguing in the context of Hck’s potential involvement in signaling pathways regulating RAG expression, since both of these Hck-associated proteins may themselves be involved in RAG regulation [49 a novel, 220]. In the former case, temperature-dependent activation of an oncogenic version of c-abl (v-abl) in pre-B cell variants was found to downregulate RAG mRNAs. Therefore, unlike hBRAG, v-abl appears involved in a negative RAG regulatory pathway [220]. Since the phosphorylation and subsequent activation of Hck has been reported to prevent its association with Bcr-abl [418], it would be interesting to test if Hck may indirectly act as an indirect positive RAG regulator via negative regulation of the potential negative RAG regulator c-abl. With respect to Hck’s association with hBRAG, hBRAG may in turn indirectly regulate RAG expression via Hck regulation. In this context, α-phosphotyrosine immunoblots of Hck immunoprecipitates suggest that Hck phosphorylation is increased upon BCR ligation (Figure 4.7B), a finding that is consistent with a previously report of BCR-ligation mediated increases in Hck phosphorylation/activation [415]. In contrast, the hBRAG-associated Hck fraction appears to be dephosphorylated upon BCR stimulation (Figure 4.7B), suggesting the possible differential regulation of Hck by hBRAG. One mechanism for how hBRAG may therefore mediate a positive effect on RAG expression via Hck regulation could be by negatively regulating Hck activity, the latter which when not associated with c-abl may act as a negative, rather than positive RAG regulator.
Chapter 4

Overall, it is tempting to speculate that Hck may be involved in a positive RAG regulatory pathway, either by association with hBRAG or by dissociation with other proteins such as Bcr-abl. Hck also appears to be involved in multiple other hematopoietic cell signaling pathways, including possibly cytokine and B cell co-receptor signaling pathways, and thus may also be involved in a B cell-specific RAG regulatory pathway in this context (reviewed in [19, 20]).

4.4.6 Anti-hBRAG ligation alone generates modest increases in protein tyrosine phosphorylation and may also increase the degree of BCR-mediated phosphorylation upon its co-ligation with anti-μ.

The fact that hBRAG potentially associates with molecules that are phosphorylated in resting B cells, and are altered in this process upon BCR ligation, coupled with its potential association with Hck, and the structural clues presented in this and the previous chapter, all suggest that hBRAG is capable of functioning as a cell surface receptor with signaling capabilities. Although a large portion of the glycosylated and unglycosylated forms of this protein seems to be intracellularly retained, the subcellular localization data is not inconsistent with a signaling capability as a low level can also be expressed on the cell surface. To assess hBRAG signaling potential in the context of the B cell lineage, the putative hBRAG cell surface receptor was ligated with α-hBRAG antibodies and proximal signaling parameters i.e. phosphorylation patterns in whole cell lysates were assessed using an α-phosphotyrosine antibody in immunoblots (Figure 4.8A). α-hBRAG stimulation at presumably saturating concentrations (10 μg/ml) generated a low, but detectable tyrosine phosphorylation response relative to unstimulated cells and to the irrelevant rabbit IgG control, but one that was much lower than via BCR stimulation. It is difficult to ascertain from these experiments if the phosphorylation pattern is exactly the same or partially overlapping, but distinct from that of BCR ligation (Figure 4.8A). Interestingly, when hBRAG was ligated in combination with BCR, this increased the intensity and the pattern, of BCR mediated tyrosine phosphorylation, suggesting it may positively enhance early activation events through the BCR (Figure 4.8A, lanes 8 and 9). However, unexpectedly, ligation of the BCR in combination with the irrelevant IgG control also increased BCR-mediated phosphorylation almost to the same degree as α-hBRAG+anti-μ co-ligation (Figure 4.8A, lane 10). One possibility for this effect is that the
**Figure 4.8. Analysis of tyrosine phosphorylation profiles in A8-6P whole cell lysates cross-linked with anti-hBRAG, anti-μ, or both.**

A. Relative contribution of hBRAG in protein tyrosine phosphorylation induction in B cells. Whole cell lysates were prepared from A8-6P endogenous hBRAG-expressing mature OCI LY8-C3P variants that were either unstimulated (lane 1), stimulated with a goat F(ab)'2 polyclonal α-human μ antibody fragment alone (lanes 2, 3), hBRAG antibodies alone (affinity purified from terminal bleed antisera raised against the N-terminal peptide UT952; lanes 4-6), or ligated with both (lanes 7-9). The durations and concentrations of antibody stimulation are indicated in each lane. To control for non-specific effects, ligation with 10 μg/ml of irrelevant affinity purified rabbit IgG was used in combination with 10 mg/ml anti-μ for 3 minutes (lane 10).

B. Relative effect of specific hBRAG ligation in enhancing BCR-mediated tyrosine phosphorylation. Whole cell lysates were prepared from A8-6P endogenous hBRAG-expressing mature OCI LY8-C3P variants that were either left unstimulated, or stimulated for 3 minutes with either 0.1 or 10 μg/ml of α-human μ alone (lanes 2 and 3, respectively), with 10 μg/ml of both α-human μ and α-hBRAG (lanes 4-6) or the combination of 10 μg/ml of α-human μ and irrelevant IgG control. To test for the relative contribution by Fc receptor components, A8-6P cells were preincubated with 10μg/ml blocking antibodies directed against human IgG Fc receptors (+Fc block, lanes 5 and 8, respectively). To test for specificity of the hBRAG mediated effect, hBRAG antibodies were pre-incubated with 10μg/ml of the competing peptide UT952 (+UT952, lane 6). In both panels A and B, 2x10^7 cell equivalents were fractionated by SDS-PAGE run under denaturing conditions, transferred to a PVDF membrane, and phosphotyrosine-containing proteins were detected using the 4G10 antibody. Blots were also stripped and re-probed with an α-actin antibody to control for loading differences.
intact IgGs in these studies may non-specifically alter BCR signaling by co-crosslinking of F\(_2\gamma R\)II/III receptors. F\(_e\) receptor ligation is generally thought to be involved in negative signaling via activation of their ITIM motifs and association with shp1 SH2 domains resulting in the subsequent dephosphorylation of selected downstream signaling intermediates [128]. In the context of B cells, F\(_2\gamma R\)II co-ligation has been shown to inhibit several consequences of BCR stimulation, including phosphoinositide hydrolysis, intracellular calcium flux, cellular proliferation, and Ig secretion. However, co-ligation of F\(_2\gamma R\)II with BCR i.e. ligation with intact anti-\(\mu\) and its effect on resulting phosphorylation patterns, although not well documented in human mature B cells, appear to produce either similar or more intense overall tyrosine phosphorylation patterns in whole cell lysates of mouse BCR\(^*\) B cells and B cell lines relative to F(ab)'\(_2\) anti-\(\mu\) cross-linking (reviewed in [128]).

To test for F\(_e\)-mediated non-specific effects and specific hBRAG effects in combination with BCR ligation, the effect of blocking specific and non-specific components of hBRAG ligation was tested by pre-incubation of \(\alpha\)-hBRAG antibodies with saturating concentrations of UT952 peptide and pre-incubation of A8-6P cells with saturating concentrations of blocking antibodies directed against the F\(_2\gamma R\)II/III receptor, respectively. The results of these experiments suggest that increases in tyrosine phosphorylation in A8-6P cells are at least partially mediated by non-specific F\(_e\)-mediated effects, as a decrease in phosphorylation was observed with IgGs or hBRAG ligation in the presence of the F\(_e\)-blocking antibody (Figure 4.8B, lanes 2 and 4, respectively). However, a component of the \(\alpha\)-hBRAG ligation effect also appears to be specific as blocking with UT952 has a substantial effect in suppressing the enhanced BCR-mediated phosphorylation pattern (Figure 4.8B, lane 6). The various differences in signal intensity were not due to loading differences, as demonstrated by re-probing immunoblots with an \(\alpha\)-actin antibody. To test more definitively for the specificity of hBRAG alone and/or in conjunction with BCR signaling in inducing tyrosine phosphorylation of downstream proteins, it will be important to generate an F(ab)'\(_2\) hBRAG fragment, and to test the kinetics of these effects in normal human B cells and other BCR\(^*\) B cell lines.
Overall, based on the results from co-immunoprecipitation, subcellular localization and signaling experiments, coupled with hBRAG mRNA expression correlating with B cell-specific RAG expression, we hypothesize that hBRAG mediates potentially important regulatory signals in a B cell-specific signaling pathway. Such a pathway may be important for B cell development and/or regulating B cell-specific RAG expression. In doing so, it may be involved in “negative” signaling pathways, but how this fits in with positive RAG regulation is not clear based on what is reported in the current literature. Various signaling molecules, however, have functions in RAG regulation at various stages of B cell development such as CD19, CD40, IL7R and the BCR. Some of these, such as CD19, in addition to positive signaling functions, have also been shown to have roles in “negative” signaling [122]. CD19 can lower the threshold of BCR ligation when co-ligated in conjunction with BCR but conversely can render a cell refractive to BCR signaling when pre- ligated alone [115]. In this context, it would be interesting to ascertain if hBRAG may function in an analogous way to CD19, in that it may exhibit biphasic signaling responses, and may be involved in an ITIM-independent negative signaling pathway. Furthermore, hBRAG like CD19, may be an example of a molecule that can act both in altering BCR-mediated early activation events as well as having a role in altering RAG expression in human B cell development.

Although overall results are consistent with hBRAG functioning as a cell-surface signaling receptor molecule, it is possible that hBRAG may also have a function as an intracellular enzyme. The large intracellular fraction of hBRAG seen in Figure 5B would support a role for hBRAG distinct from that as a cell-surface signaling receptor. Further experiments, including indirect immunofluorescence, subcellular fractionation, and molecular characterization of hBRAG-associated proteins should help to more precisely define the function of hBRAG protein isoforms in different subcellular locations.
CHAPTER 5:

General Discussion
5.1 Technical considerations and refinements in screening for RAG co-expressing cDNAs by DD RT-PCR

5.1.1 Pros and cons of DD RT-PCR relative to other techniques for assaying differential gene expression

In techniques such as subtractive hybridization, cDNA-RDA, or microarray analysis, only one cell population at a time can be feasibly studied. One important advantage of DD RT-PCR over other techniques is its ability to simultaneously compare (display) genes that are differentially expressed from multiple sources [419, 420]. This opens the door for various possibilities, for example, stronger correlations of candidate genes can be made if samples from multiple patients with the same disease are compared, different timepoints of induction can be assessed, and multiple developmental stages can be assayed systematically. Therefore, testing several independent sources simultaneously not only increases the reproducibility of the assay, but potentially allows for more physiologically relevant genes to be isolated and tested functionally. Another important advantage DD RT-PCR has over most techniques, especially earlier methodologies such as subtractive hybridization and differential screening, is that only miniscule amounts of total RNA are required [314]. This is particularly advantageous for identifying developmental genes because often these are only found in discrete cell subsets from which only small amounts of total RNA can be obtained (for example, developmental lymphocyte subsets sorted by flow cytometry) [368]. DD RT-PCR is also advantageous to other techniques (with the exception of SAGE, see Figure 1.15) in that it reveals both upregulation and downregulation as well as subtle quantitative differences in gene expression. The latter is an important consideration in studying molecular programs of lymphocyte development because lymphocyte "stage-specific" genes may not be completely turned off (for example, the RAG genes) and such subtle differences in gene expression may translate into functional differences (in the case of the RAG genes, recombinase activity). Finally, the speed and relative-straight forward methodology in which candidate genes can be isolated by DD RT-PCR compared to other techniques (such as subtractive hybridization and differential screening) makes this method an attractive alternative.

The most obvious disadvantage of DD RT-PCR and other comparative techniques for identifying differentially expressed genes relative to subtraction-based techniques (such as subtractive
hybridization or cDNA-RDA) is that one is not enriching for differences, with the result that in many more candidate genes having to be subsequently screened. Relative to non-PCR based techniques, DD RT-PCR has several drawbacks. The first is the preponderant inability to reproduce the pattern of differential expression seen on the display in Northern blots i.e. generation of false positive and false negative artifacts. In fact, most investigators currently using the technique are reporting an inability to reconfirm differential expression at a rate of $\geq 85\%$ [314, 421]. False negatives pose a major problem, as a large number of rare message class genes, including developmentally associated factors, will be potentially missed using this technique. Another problem that has recently come to light with the DD RT-PCR technique is the heterogeneity of DD amplicates i.e. the presence of multiple amplicons actually comprising the reamplified, differentially expressed DD amplicon [314]. Amplicate heterogeneity is best indicated by the experience of many investigators, where multiple transcripts are detected on Northern blots when uncloned DD amplicons are used as probes [314]. Impure mixes of products necessitate further laborious rounds of screening and purification, thus making the procedure impractical on a large scale. What are the reasons for the aforementioned limitations? There are multiple sources for each limitation stemming from various aspects of the procedure. These are outlined in Table 5.1 and will also be discussed below.

5.1.2 An in-depth look at the limitations of DD RT-PCR and current proposed refinements/modifications to the technique

Both amplicate heterogeneity and false positives can arise from the differential display procedure itself. In particular, DD RT-PCR amplicate heterogeneity may result due to low-stringency of the initial RT-PCR amplification step. This may allow the amplification of potentially large numbers of co-migrating, high abundance, contaminating non-specific amplicates that “mask” bona fide differentially expressed amplicons on DD RT-PCR gels (resulting in false positives) or to preferential amplification of amplicons within the same amplicate (resulting in amplicate heterogeneity). However, the high rate of heterogeneous amplicates/false positives is usually not due to the differential display procedure itself, as long as appropriate controls are run i.e. use of independent RT and PCR reaction duplicates and multiple DNA-free total RNA samples. The major sources of false positives and amplicate heterogeneity originates from the steps subsequent to the display, namely the excision, reamplification,
Table 5.1
Summary of limitations to analysis/selection of DD RT-PCR cDNA products and suggested refinements to original DD RT-PCR and subsequent screening steps.

<table>
<thead>
<tr>
<th>Technical Limitation</th>
<th>Reason for limitation</th>
<th>Stage of DD RT-PCR or subsequent screening procedure affected</th>
<th>Potential problematic source</th>
<th>Suggested improvement, refinement/reference(s)</th>
<th>Intended effect of improvement</th>
<th>Potential drawbacks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laborious, Inefficient screening</td>
<td>High % false positives i.e. no differential expression seen by Northern blot</td>
<td>PCR amplification</td>
<td>Mispriming under low-stringency conditions</td>
<td>duplicate PCR and RT reactions (Liang et al., 1993; [311])</td>
<td>Increased reproducibility in cDNA banding pattern on DD gel</td>
<td>increased labor</td>
</tr>
<tr>
<td>Display of DD amplicons</td>
<td>identical sized co-migrating, high abundance, non-specific cDNAs on sequencing gel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excision of DD amplicons from sequencing gels</td>
<td>Inaccuracy of excision process; &quot;bunching&quot; of DD bands at top of gel</td>
<td>Longer gel runs (Liang and Pardee, 1995; [314])</td>
<td>more precise excision</td>
<td>increased labor (still need short gel for lower products)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR reamplification</td>
<td>Mispriming under low-stringency conditions</td>
<td>Reamplification with the same 5' arbitrary decamer and each individual 3' anchored primer (Sompayrac et al., 1995; [421])</td>
<td>Eliminates mispriming in which either the 5' or 3' primers act alone as primers for both ends</td>
<td>only eliminates a particular class of mispriming, increased labor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern blot screening</td>
<td>Low stringency of washings, resulting in masking of differentially expressed transcripts by co-migrating non-specific transcripts</td>
<td>Increasing washing stringencies (Verkoczy and Berenstein, 1998; [392])</td>
<td>Elimination of masking, high abundance, co-migrating transcripts?</td>
<td>also may eliminate detection of differentially expressed transcripts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excision of DD amplicons from sequencing gels</td>
<td>Missing DD amplicons in excision process</td>
<td>Longer gel runs (Liang et al., 1994; [419])</td>
<td>more precise excision</td>
<td>increased labor (still need short gel for lower products)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern blot screening</td>
<td>incapability to detect rare message class differentially expressed genes</td>
<td>RNAse protection assay as alternative to northern blotting (Yeatman et al., 1995; [422])</td>
<td>Increased sensitivity-differentially expressed, rare-message transcripts may be detected</td>
<td>limited practicality for large # samples being screened for</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
- Potential problematic source: Problems observed in the analysis/selection of DD RT-PCR cDNA products.
- Suggested improvement, refinement/reference(s): Proposed solutions to improve the analysis/selection process.
- Intended effect of improvement: Expected outcome of the suggested improvements.
- Potential drawbacks: Possible limitations or challenges associated with the proposed solutions.
and Northern blotting procedures. For example, false positives can result from a differentially expressed product may be masked on Northern blots by closely or identically co-migrating higher abundance transcripts. A second major source for false positives post-DD RT-PCR is due to the inability to accurately resolve or "display" banding patterns (overlapping bands often occur in the 400-600 bp range where many of the desired amplicons migrate and where the pattern of banding on the sequencing gel is highly compressed), improper excision of amplicates in these portions of the DD RT-PCR gels may also generate false positives. With respect to amplicon heterogeneity after the display step, this can be due to the PCR reamplification (especially when two rounds of reamplification are required), which can cause contaminating amplicons not seen in the original display to become overrepresented or can be due to the radiolabeling step, where higher abundance amplicons may outcompete bona fide differentially expressed amplicons.

False negatives arise due to most of the aforementioned reasons responsible for generating false positives and amplicate heterogeneity. However, because differentially expressed transcripts are often of the rare message class [314], they are often the result of a source that uniquely generates false negatives: insufficient sensitivity at the Northern blotting step. One way to deal with the low sensitivity of Northern screening can be to use more sensitive RNA quantitation methods. For example, a modified RNAse protection assay has been described [422]. In this assay, primers used in the 2nd round reamplification step are designed in such a way that the DD RT-PCR products can be directly used in RNAse protection without prior subcloning steps.

In the DD RT-PCR screens described in this thesis, reamplification and excision issues were particularly relevant, because conditions that were optimized for the detection of rare message class transcripts were also found to be very promiscuous, as indicated by several observations. First, internal sequences were frequently amplified instead of the expected 3' poly-adenylated ends. Secondly, in our secondary and tertiary screenings, 3/4 of the original amplicons were found to be non-specific. Thirdly, amplicons that were reamplified were actually a mixture of amplicons where several high-abundance, non-specific transcripts predominated.
Since the demands of the experiment make it hard to change conditions for the initial amplification, several promising improvements for decreasing sources that potentially generate both false positives/negatives at stages either prior to or subsequent to the DD RT-PCR amplification step have been proposed. For example, a preliminary round of subtraction can be added to the procedure prior to any DD RT-PCR steps. A subtraction step could potentially eliminate a large number of high abundance transcripts, which can lead to both false negatives and positives [423]. Another proposed refinement has been to insert SSCP analysis of excised DD products prior to the reamplification step as another screen for eliminating co-migrating, high abundance, non-specific contaminants [424, 425]. Longer gel runs may also add increased accuracy to the procedure. However, appealing though these refinement may be for large-scale screens, it cannot circumvent the limitation of amplicate heterogeneity. Regardless of various refinements when heterogeneity of amplicates exists, the isolation of the desired amplicon is at best a "hit or miss" proposition. It is impossible to know in advance which clone contains the amplicon corresponding to the differentially expressed transcript. Furthermore, if cloning efficiency is low, the underrepresented, "correct" clone may not be obtainable. Additionally, these steps significantly increase the time required for the overall screening procedure, which is one of the attractive advantages of DD RT-PCR relative to other techniques in the first place.

Modifications of the reamplification step have also been proposed, such as increasing its stringency [426], but this may eliminate the differentially expressed transcript altogether especially if it becomes underrepresented relative to other contaminants. Two recent reports describe methods that show promise for dealing with amplicate heterogeneity subsequent to the procedure's reamplification phase. In one report, "affinity capture" is described, where the radiolabeled cDNA from the DD amplicon probe is eluted directly from the Northern blot and then subjected to reamplification and purification [374]. This has also been an approach that I have used with success in my studies (Chapter 2, Figure 2.6 and [392]). Another suggested possibility is to combine DD RT-PCR with a subsequent differential screening procedure [377]. In this assay, candidate DD amplicates are directly subcloned as sublibraries, and colony lifts are probed with labeled cDNAs from the two samples being compared.
Technical limitations leading to amplicate heterogeneity, the high rate of false positives and/or false negatives are not the only factors that make DD RT-PCR an inefficient procedure. The fact that DD RT-PCR is a comparative technique, also makes it a very laborious approach when applied on a large-scale i.e. in the comparison of expression profiles of all genes between several sources. For example, it is not uncommon for a DD RT-PCR screen to yield >100 candidate clones, a daunting number for further screening and selection purposes. In light of this, several ways to streamline the overall screening procedure have been proposed (see table 5.1 under theoretical considerations). Perhaps one of the more promising of these is to use a "reverse dot blotting" procedure, in which DD clones are dot blotted onto membranes and probed in parallel with radiolabeled cDNAs derived from the samples being compared [375, 427].

5.2 General structural and functional implications of hBRAG studies

5.2.1 The potential role of hBRAG in normal B cell development, RAG regulation, and V(D)J recombination: general comments

The hBRAG mRNA expression analysis and transfection studies suggest that hBRAG may be important in B cell differentiation and as such, may regulate RAG specifically in the context of B cell differentiation. Thus hBRAG may also be an important control molecule in the process of Ig rearrangements. The role of hBRAG in B cell-specific regulation of RAG mRNAs will be discussed later. Furthermore, analysis of hBRAG mRNA expression in various human tissues (where it was found to be highest in PBL and spleen tissues) would suggest it may have a role in mature B cell development. The high levels of expression of this potential RAG-regulating gene in mature B cells would be consistent with RAG re-expression in GCs. Whereas anti-Ig signaling is involved in downregulation at this stage, it is therefore possible that hBRAG is involved in the induction of this "fourth wave" of RAG expression.

Based on the biochemical data presented in Chapter 4 and on structural and biochemical similarities and aa sequence homologies to other proteins, hBRAG can be predicted to affect B cell development and RAG regulation by several possible mechanisms. As a cell-surface protein, it may have a direct function in these processes either as a cell surface signaling receptor or an ectoenzyme. As an
intracellular protein, it may have a more indirect effect, in particular as either an enzyme or chaperone protein involved in the modification/regulation/transport of RAG-regulating and/or B cell-specific molecules. Its potential role in each of these contexts will be discussed further below.

5.2.2 Structural comparison of hBRAG to other characterized type II proteins in the immune system

hBRAG has several unique structural features distinguishing it from other characterized type II TM proteins that function in lymphocyte regulation/development and/or with lymphocyte-specific expression patterns (discussed in section 1.5 of the General Introduction). First, hBRAG has no overall homology to any of these proteins. Secondly, it is likely not a multigene family member based on genomic southern blot analysis (See Chapter 3, Figure 3.7). Thirdly, unlike most type II TM molecules with immunologic properties which are generally expressed in either T, NK, or monocyte lineages, hBRAG is preferentially expressed in B cells, and at low or undetectable levels in T cells (Chapter 3, Figures 3.5 and 3.6). The only other type II proteins that are preferentially expressed in B cells are CD23 and CD72. Unlike these two proteins, hBRAG lacks CRD domains in its extracellular region, and has no cytoplasmic inhibitory ITIM motif. Fourthly, hBRAG has a shorter TM than the above-mentioned cell-surface molecules, a property that is associated with intracellularly expressed, golgi-associated integral membrane proteins (reviewed in [401]). The fact that the in vitro translated hBRAG product in the presence of microsomes is efficiently targeted to the Golgi membrane and glycosylated and the relatively large fraction of intracellular hBRAG in biotinylation experiments both reinforce this notion (See chapter 4, Figures 4.2 and 4.6B). The invariant chain is the only known type II TM protein with characterized immunologic function that is predominantly retained intracellularly. Its intracellular localization, however, can occur via either its retention or endosomal-targeting signals, neither one which hBRAG contains in its predicted structure. Like TNF ligands and scavenger receptors, hBRAG is rich in cysteine-rich regions, but are not similar enough to constitute a TNF or SRCR motif. hBRAG is also more extensively glycosylated (at least in vitro) than either TNF or C-lectin receptors (see Chapter 4, Figure 4.2), which only have one or two N-glycosylation sites/polypeptide monomer (the only exception to this being the CD30L).
Some generalities can be made about characterized type II TM proteins in the immune system that may reinforce and/or yield further clues about hBRAG structure/function. First, with the exception of invariant chain, they are all preferentially expressed on the cell-surface. Even invariant chain, which is a predominantly intracellularly-expressed protein, without its ER retention signal, can also be expressed on the surface at low, but detectable levels. The low, but detectable levels of hBRAG in cell surface staining and biotinylation assays (Chapter 4, Figure 4.6) is therefore consistent with the structural prediction that hBRAG, which lacks such a signal, should at least partly leak out to the plasma membrane. Secondly, many type II TM proteins exist function as higher order structures on the cell surface (Figure 1.18); either as homodimers (CD69, CD72, LY49, CD94), heterodimers (CD94, NKG2) or homotrimers (CD23, Fas-L); invariant chain also exists as a homotrimer (or as a nonamer with class II association). Consistent with hBRAG existing as a homo- or heterodimer, higher molecular weight isoforms are also seen in α-hBRAG immunoblots of whole cell lysates or of α-hBRAG immunoprecipitations of de novo synthesized proteins run under non-denaturing conditions (see Chapter 4, Figures 4.3 and 4.4 and results/discussion section).

5.2.3 Speculations as to molecular mechanisms of hBRAG function in normal B cell development and RAG regulation

Cell surface hBRAG as a signaling receptor

All the cell surface proteins described above (excepting BP-1; see below) with lymphocyte-specific function and/or expression patterns have characterized signaling potential, whether negative i.e. the C-lectin family or positive/co-stimulatory i.e. the TNF family mostly involved in T-T and T-B cell interactions. If one extends this observation to all lymphocyte-specific type II proteins, there are several structural and functional features about the cell surface-expressed hBRAG fraction that supports its role as such a signaling receptor. First, it has candidate signaling motifs in its 81 aa cytoplasmic N-terminus including consensus cAMP and PKC phosphorylation sites and palmitylation residues (See Chapter 3, Figure 3.4), shared by various other signaling receptors and intermediates, such as G protein-linked receptors, G proteins, and Src family kinases including Src, Blk, Lyn, Yes, Hck, Fyn, and Lck (reviewed in [410]). Additionally, in a functional context, α-hBRAG co-
immunoprecipitation results suggest that it and/or various phosphorylated molecules it associates with are selectively dephosphorylated (and/or dissociated) within one minute post-BCR ligation (See Chapter 4, figure 4.5). This suggests that hBRAG either directly phosphorylates these molecules or associates with molecules that are phosphorylated in response to BCR cross-linking. Finally, α-hBRAG treatment of cell produces low, but detectable increases in tyrosine phosphorylation relative to controls, and may also enhance the BCR-mediated early activation events (Chapter 4, Figure 4.7). Assuming hBRAG’s involvement in signaling, the fact that hBRAG’s sequence does not have ITIM motifs suggests its role either in positive (co-stimulatory) or shp1 independent-negative signaling.

What role would hBRAG have as a signaling molecule in B cell development and RAG regulation? hBRAG may be involved in signaling-mediated regulation of RAG expression in earlier B cell differentiation, analogous to the pre-BCR, pre-TCR, CD19, and IL-7R (see Table 1.3 and Figure 1.4 for a complete list of positive regulators of RAG expression in B cell development). Additionally, since hBRAG appears to be enriched in peripheral tissues, it may potentially deliver a signal which reinduces RAG expression and secondary rearrangements much like signaling through CD40, the IL-4R, or IL-7R in mature GC B cells (reviewed in [428]). In this regard, it could be a co-stimulatory receptor acting in conjunction with BCR signaling or alternatively, by analogy to CD40, one that acts independently of BCR ligation in mature B cells.

In the latter case, hBRAG may have signaling capabilities independent of a native ligand because it can upregulate RAG1 mRNA levels by itself, at least in the context of stable transfection studies (see Chapter 3, Figure 3.9). Alternatively, it could be a B cell-specific ligand/counter-receptor for a receptor involved in RAG induction or B cell signaling, either novel or already characterized, such as CD40, IL-7R or IL-4R. In the context of hBRAG as a possible B cell and/or RAG-specific counter-receptor, it is relevant that almost all TNF family type II TM proteins are counter-receptors for co-stimulatory receptors. Furthermore, a stromal-cell counter-receptor has been cloned that delivers a RAG-inducing signal to the human pro-B cell line PA6 [217]. One caveat, based on sequencing discrepancies between KIAA0598 and hBRAG (discussed in the Chapter 4 results and discussion and in more detail below), is that it is also possible that I have cloned a dominant "active" form of the
putative receptor. This activated form would act much like the activated \textit{lk} transgene which can activate T cell signaling independent of TCR signals [92].

\textit{Cell surface hBRAG as an ectoenzyme}

As an alternative to being a cell surface receptor, hBRAG, as a protein expressed on the surface of B cells (Chapter 4, Figure 4.6), may serve similar functions as those postulated for the type II TM ectoenzyme BP-1 (discussed in the General introduction section on type II proteins). In this context, hBRAG may either cleave products in mature B cells (where it is highly expressed) that negatively signal RAG downregulation (such as the BCR components in GCs). Alternatively, it may interact with various ligands on T cells or other APCs with substrate specificity analogous to that of the IL-7R N-terminal glutamic acid residues accessible for BP-1 modification. This type of regulatory interaction could therefore alter interactions with B cell-specific receptor components (including the Ig-\(\alpha\), Ig-\(\beta\), and \(\mu\) components of the BCR itself) and/or other positive (co-stimulatory)/negative signaling receptors mediating RAG regulation and T-B interactions in general. hBRAG as an intracellular enzyme may also be directly involved in the post-translational modification of the ligand binding specificity of another B cell-specific signaling molecule. In this context, the ligand-binding site of CD22, which has been mapped to its two N-terminal Ig domains, has its binding specificity mediated by \(\alpha\)-2,6-linked sialic acid side-chains on N-linked oligosaccharides of glycosylated proteins. One of the enzymes that effectuates the interaction of CD22 with its ligands via catalyzing sialic acid \(\alpha\)-2,6-linkage on penultimate CD22 galactose residues is \(\beta\)-galactose \(\alpha\)-2,6-sialytransferase.

\textit{Intracellular hBRAG as a chaperone molecule}

Invariant chain (Ia), as the only known intracellularly-expressed type II TM protein with characterized immunologic function, is believed to work in several ways in the transport and assembly of MHC class II molecules. As discussed in section 1.5 of the General Introduction, two mechanisms by which Ia mediates MHC class II trafficking/assembly is as a chaperone protein and as a competitive inhibitor molecule. Because the surface labeling experiments in Chapter 4 would suggest hBRAG’s preferential intracellular expression, as a golgi-associated protein, it could therefore be potentially
involved in a mechanistically analogous fashion to that characterized for Ia, in the context of the transport and assembly of B cell specific and/or RAG-regulating molecules. For example, because of its preferential expression in the B cell lineage and/or co-expression with RAGs, as a chaperone for certain ubiquitously-expressed cell surface-expressed proteins, hBRAG may allow their transport/assembly only at the highest efficiency in the context of the B cell lineage and/or at stages when RAG is co-expressed with hBRAG. Furthermore, hBRAG could be involved as a regulator in the assembly, folding, and/or intracellular trafficking/transport of such proteins (including movement from the ER, through the golgi, or targeting to various organelles such as endosomes for proteolytic cleavage). In particular, hBRAG may also act as a competitive inhibitor by blocking sites on such proteins from premature processing and/or by preventing targeting of such proteins to endosomal compartments. In the context of the role of hBRAG in the regulation of RAG regulation, by regulating the trafficking and assembly of a RAG-regulating protein, this would be an example where its effect on this process would be indirectly rather than direct, a possibility alluded to earlier in the Chapter 3 discussion.

5.2.4 Proteins with homologies to hBRAG

Although no proteins with characterized immune functions have any sequence homology to hBRAG, two other proteins, KIA0558 and heparan sulfotransferase, share at least partial sequence identities to hBRAG. These will be discussed briefly in the context of any potential further clues as to hBRAG function.

KIAA0598 protein

Recently, a protein nearly identical to hBRAG, designated KIAA0598, has been recently cloned from a brain cDNA library [409]. KIAA0598 has 99% overall cDNA identity, and 91% aa identity to hBRAG. However, since the study in question was a mass-sequencing effort, no characterization of this sequence was reported. A comparison of the KIAA0598 and hBRAG complete cDNAs and proteins are shown in Figures A.1 and Figure A.2, respectively. Overall, the cDNAs are >99% identical except that the KIAA0598 cDNA has an extra 360 bp of 5' UTR. Note that the 3' UTR
region is identical; this could be either because the same EST was used or there is no 3' UTR diversity in this gene across tissues. Within the region of >99% homology, there are three significant single bp additions or deletions in the ORFs which result in two major differences in the proteins. First, there is a nucleotide addition in the KIAA0598 cDNA corresponding to hBRAG bp 747 which results in an ORF shift in a region corresponding to hBRAG aa 180-220; a single bp deletion restores the original ORF (Figure A.1). Second, there is a single bp deletion of KIAA0598 cDNA corresponding to hBRAG bp 1671. This results in an ORF shift that replaces the stop codon at hBRAG aa 503 and results in an additional 60 aa of C-terminal sequence in the putative 563 aa KIA558 polypeptide (Appendix, Figures A.1 and A.2). Therefore, while the cytoplasmic and TM regions of the hBRAG and KIAA0598 proteins are identical, the aforementioned differences are in the hBRAG extracellular domain.

In context of the above-mentioned sequencing differences in the hBRAG and KIAA0598 3' end coding regions, the affinity purified antibodies raised against the hBRAG C-terminus failed to specifically recognize either the unglycosylated hBRAG doublet or the higher molecular weight glycosylated product (Appendix, Figure A.3, lanes 10-13). This finding is consistent with either single or dual peptide competition assays, in which the C-terminal antibody does not recognize the same proteins and/or hBRAG isoforms as the N-terminal antibody (Appendix, Figure A.4). Taken together, these results suggest that the C-terminal antibody does not recognize the hBRAG product from the predicted sequence of the 8-3 hBRAG cDNA that I cloned. It is not known if the species detected by the C-terminal antibodies correspond to isoforms of this protein, since no characterization aside from the sequence was reported in this study. The C-terminal discrepancy is in the region of the peptide used to generate the hBRAG C-terminal antibody (a detailed comparison of hBRAG with this protein is discussed further in Chapter 5 and shown in Figures A.3 and A.4).

It is not likely that the differences in hBRAG and protein sequences and/or hBRAG antibody specificity are due to sequencing errors, since in both studies, several independent cDNA clones were sequenced and verified against each other. It is formally possible, however, that such differences are due to inherent in the templates used for in vitro transcription and/or for cDNA cloning. More
interestingly, these differences could be due to functionally important allelic polymorphisms. Alternatively, this discrepancy could represent mutational differences in the hBRAG or the KIAA0598 proteins, a possibility that is consistent with the observation that transfection of hBRAG, as a putative cell surface receptor cloned from a human pre-B cell line, can have a ligand-independent effect on RAG expression. Assuming bona fide polymorphic/mutational differences in these hBRAG and KIAA0598 extracellular regions (either in the ~40 aa divergent region and/or in the additional 60 aa C-terminal KIA0598 sequence), there are no clues indicating possible functional differences, including any known motifs. In this regard, it will be interesting to test for differential effects of the hBRAG and KIAA0598 complete cDNA clones in various contexts once a better idea of gene function is ascertained.

**Heparin sulfotransferase**

Heparin sulfotransferase is a ubiquitously-expressed intracellular Golgi-associated enzyme with both \(N\)-deacetylase activity and \(N\)-sulfotransferase activit. Sulfotransferases are involved in golgi-associated biosynthesis of heparin and heparin sulfate, two molecules with diverse functions including cell recognition and adhesion \([431, 432]\). Heparan sulfotransferase has several structural features suggesting its evolutionarily relationship to hBRAG. First, its locus is found in relatively close proximity to that of hBRAG on human chromosome 10q26 (see Chapter 3 discussion). Second, it is also a type II TM glycoprotein that shares restricted homology to hBRAG in the extracellular domain (~30% over a 200 aa stretch). Finally, like hBRAG, it has a short TM, and is integrated in the microsomal fraction at high efficiency *in vitro*, consistent with its intracellular expression.

In the context of possible structural and functional homologies with the heparan sulfotransferase protein, the large intracellular hBRAG fraction may be enzymatically involved in the post-translational biosynthesis of B cell-specific and/or RAG-regulating proteoglycans. Since many B cell-specific signaling receptors and intermediates are heavily glycosylated (reviewed in \([417]\)), they may require a protein such as hBRAG to specifically post-translationally modify them for their proper functioning. Alternatively, as a type II golgi-associated enzyme, hBRAG may be involved in post-translational
control of B cell-specific signaling intermediates and/or transcription factors. A large number of type II TM proteins which are Golgi-associated (which the majority of type II proteins are, in general) have a role in post-translational modification reactions that may be involved in functional regulation [339]. Post-translational regulation of proteins is an important level of regulation that can be effected in various ways including phosphorylation, dephosphorylation, sulfonation, and glycosylation, as just several examples [433].

5.2.5 hBRAG as a Rag-associated protein?

Another less likely interpretation of hBRAG and RAG co-expression in B cell lines is that hBRAG might perform functions similar to RAG in V(D)J recombination, and may thus increase the overall efficiency of the V(D)J recombination reaction, as a "RAG3-like enzyme". Not mutually exclusive with the above possibility, hBRAG could directly interact and modify the Rags allosterically. In this context, two proteins, Rch1 and SRP1, have been characterized as directly interacting with Rag1 via the N and C-terminal regions of Rag1, respectively [165 a protein, 166]. Additionally, with the characterization of Rag cleavage and post-cleavage complexes, a whole other series of proteins are likely to either directly or indirectly interact with the Rags, including HMG1/2, the Ku heterodimer, and the DNA-dependent protein kinase catalytic component (DNA-PKcs) [167, 434]. As part of these complexes, hBRAG could therefore have a role in cleavage or post-cleavage events. The possibility of RAG either being a recombinase component and/or a RAG-associated protein at present seems unlikely because the hBRAG aa sequence does not have a NLS. Also, nuclear-associated proteins in general do not have transmembrane regions. It is formally possible however, that it may be nuclear-associated indirectly via an interaction with RAG1 in an as yet unidentified domain or via another protein that forms part of the cleavage or post-cleavage complex.
5.2.6 Future directions with respect to hBRAG

5.2.6.1 Further characterization of hBRAG protein and potential associations

Further biochemical characterization of hBRAG protein

To reduce the background due to cross-reactivity seen when using polyclonal hBRAG antisera, transient COS transfectants with epitope-tagged hBRAG expression constructs could be generated. This will allow detection of the hBRAG protein by FACS and immunoprecipitation with not only α-hBRAG antibodies but more specific also α-tag antibodies. Once specific detection of hBRAG has been optimized with the α-hBRAG antisera on these lines, the anti-hBRAG antisera can be used for more precise cellular localization studies. These may need to be addressed by the combination of immunohistochemistry and biochemical subfractionation, in addition to flow cytometry and surface labeling with biotin. As already discussed in Chapter 4, it will be critical to conclusively establish whether the complete tissue pattern of hBRAG protein expression is similar to the pattern of expression seen at the mRNA level. These antisera may also be potentially useful for many other studies. For example, to further study the role of hBRAG as a cell-surface expressed signaling molecule, it could be further assessed how BCR+hBRAG co-ligation cause phosphorylation, proliferative, and other signaling parameter differences relative to BCR or hBRAG ligation alone.

Identification of hBRAG-associated proteins

The above antisera have also been used to detect hBRAG-associated proteins by co-immunoprecipitation with anti-hBRAG. Based on co-immunoprecipitation experiments presented in Chapter 4, none of the known candidate B cell signaling-associated PTKs in the size range of the major dissociated/dephosphorylated product associate with hBRAG (except for low levels of the PTK Hck). An alternative would therefore be to identify the major protein(s) which are co-immunoprecipitated with α-hBRAG by microsequencing. Candidate peptide fragments derived from the associated proteins could then be entered into protein databanks for correspondences to entire cloned proteins, perhaps with known function. If such peptides have no matches, these fragments could subsequently serve as templates for generating cDNA probes by PCR and used to screen a cDNA library. The same
protein-protein interactions could be confirmed and/or other interactions could be isolated (including those with the Rags) using yeast two-hybrid assays, a technique that has been used for identifying the Rag-associating proteins SRP1 and Rch1 [165 a protein, 166].

**Generation of recombinant hBRAG protein and potential molecular and functional identification of a hBRAG counter-receptor**

To further assess hBRAG B cell-signaling potential, the experiments in Chapter 4 could be repeated with F(ab)', versions of α-hBRAG antibodies to rule out Fc-mediated effects seen in controls. Effects on other more downstream signaling parameters could be tested including [³H] thymidine uptake proliferation assays, measurement of immediate-early gene expression (such as c-fos and c-jun) or measurement of B cell activation markers (such as CD72 or CD25) by FACS. If there are strong indications that hBRAG is a co-stimulatory signaling receptor which may be involved in B-B or T-B interactions, a search for an interacting counter-receptor i.e. a natural ligand for hBRAG could be undertaken. To find the natural ligand for hBRAG, one might generate recombinant hBRAG using the 8-3 cDNA expressed in a eukaryotic expression system. This recombinant hBRAG could be used to screen a T and/or B cell receptor expression library. The recombinant hBRAG could also potentially be useful in functional assays of T-B interactions. In particular, to test for potential T-B co-stimulatory function, recombinant hBRAG would be added to T-dependent proliferation assays. Furthermore, if an hBRAG counter-receptor (ligand) is molecularly identified, one could also use antibodies raised against this ligand to test for neutralizing activity. Such interactions could also be eventually further characterized by transfection of the putative hBRAG ligand into COS cells, and testing for proliferation in this context versus that as a soluble ligand.

**5.2.6.2 Approaches towards understanding hBRAG function in RAG regulation and B cell development**

*Testing for the developmental/lineage specificity of hBRAG in the regulation of RAG expression*

In the studies presented in Chapter 3, I have stably transfected hBRAG in the context of mature B cell lines (where hBRAG was found to increase RAG1 expression), and in a non-lymphoid cell line
(where hBRAG was found to have no effect of RAG1 expression). However, I do not yet have conclusive evidence that hBRAG is a B cell-specific positive regulator of RAG expression or whether increased RAGs can be seen in other cell types, including T cells, where it is not physiologically expressed. Additionally, with respect to B cell development, it is not clear if hBRAG has a role in RAG regulation at earlier stages of development, such as in the induction of the first, second, and third "waves" of RAG expression. To further test the functional relationship between hBRAG and RAG, the converse experiment to those described in Chapter 3 would be to assess for loss of RAG1/2 expression upon introduction of hBRAG antisense constructs in the high RAG-expressing cell line C3-A11N. Furthermore, to determine whether this regulation occurs only in the context of the B cell lineage, transient transfection assays could be employed in cell lines from differing lineages and developmental stages, and with differing levels of constitutive RAG1/2 expression. Tetracycline-inducible promoters could also be used to assess the kinetics and specificity of this reaction. It may also be functionally informative to compare V(D)J cleavage activity in hBRAG-transfected cells relative to untransfected controls. This could be done by using LM PCR, an assay which detects signal end intermediates that have been generated by RAG complex cleavage in vitro [435].

In the context of the correlation between hBRAG and RAG expression in B cells, it would also be relevant to see if these two genes are co-regulated under different conditions. As discussed in Chapter 3, no effect on hBRAG expression was observed under conditions where RAG1 can be upregulated (low affinity BCR ligation), at least in the mature human B cell line OCI LY8 cell line. However, since many factors can influence RAG expression positively and negatively throughout B cell development (see Figure 1.4), hBRAG co-modulation could be tested in any of these contexts. This could include testing for hBRAG co-regulation in mature GC B cells via signals such as LPS+IL-4 or IL-7 and/or anti CD40 stimulation), and in earlier development with stimuli such as caffeine and IL-7.

**Testing the role of hBRAG in normal B cell development**

To assess what other target genes are regulated by hBRAG (if any), including other lymphocyte and/or B cell specific genes, a technique that can assess a large number of genes simultaneously is required.
Novel genes could be determined by DD RT-PCR of stable hBRAG-transfected versus vector alone-transfected OCI LY8-C3P and K562 cell lines that have already been generated. Alternatively, because a large number of human and mouse genes are currently catalogued and available as partial EST sequences, one can compare these lines using either cDNA microarray analysis and/or SAGE, both which simultaneously monitor expression levels many genes and which rely on the availability of partial cDNA information (see General Introduction sections 1.3.3.2, 1.3.4.4; Figure 1.16 and Table 1.6). A more straightforward alternative to the above approaches would be to use RT-PCR, but it will be impossible to identify either other unknown genes or characterized genes with as-yet uncharacterized lymphocyte-specific functions using this approach.

Targeted disruption of hBRAG by homologous recombination will perhaps be most important for elucidating the specificity of its function, in the immune system or elsewhere. Based on the association between hBRAG and RAG and the high expression of hBRAG in B cell-enriched tissues, it will be particularly interesting to assess the hBRAG<sup>−</sup> phenotype with respect to B cell development. Prior to generating an hBRAG-targeting construct and subsequent generation of hBRAG<sup>−</sup> ES cells, the mouse homologue cDNA and genomic clones will have to be obtained. If the targeted disruption of hBRAG results in a lethal embryonic phenotype, or has pleiotropic effects in other lineages (a possibility since hBRAG is expressed at detectable levels in other tissues), the role of hBRAG in the lymphocytic compartment could be specifically studied using RAG2 blastocyst complementation analysis [89]. In this context, it may also be useful to look at expression in embryonic development prior to attempting the knockout, since this may aid in determining whether to proceed directly to blastocyst complementation.

5.3 Identification of BSAP as a RAG co-expressed candidate in DD screens

5.3.1 The potential role of BSAP and other upstream factors in B cell-specific RAG regulation

It is interesting that the paired-domain containing transcription factor BSAP (B cell lineage-Specific Activator Protein; also known as Pax-5 or NF-HB) was identified as one of the cDNAs in my DD RT-PCR screens preferentially expressed in the high RAG-expressing variant C3-A11N. Additionally, in
further screens, I have shown that BSAP correlates with RAG1 expression in human pre-B and mature B, but not T cell lines, suggesting a potential B cell-specific regulatory role for BSAP in RAG expression. Several other indirect clues also indicate that this transcription factor which is required for B cell lineage commitment [370] and differentiation [369] may also be an important B cell-specific regulator of RAG transcription. First, in the mouse, it has been previously established that BSAP is preferentially expressed in pro-B, pre-B, and mature B cells as well as the CNS and testis [436]. Second, other B cell-specific genes critical for developmental progression in this lineage all have BSAP binding motifs in their promoters including CD19, VpreB, and λ5 [436]. Additionally, mice with targeted disruptions in the BSAP gene display a lack of pro-B cells in fetal liver, and bone marrow B cells are developmentally arrested at the Hardy fraction B stage, an earlier arrest than in RAG+/− mice, in which the block occurs at the large CD43+ pro-B cell to small CD43− pre-B cell transition (fraction C; see Figure 1.7) [369]. Finally, as discussed in the General Introduction, a recent report by Lauring and Schlissel has identified BSAP as a critical factor for B cell-specific transactivation of the RAG2 promoter [271].

Unlike RAG+/− pro-B cells, BSAP+/− pro-B cells exhibit detectable, albeit reduced levels of RAG1/2 transcripts and HC rearrangements [369]. This in vivo dispensability of BSAP in B-cell specific RAG regulation can be interpreted in several ways. First, it may be that BSAP in combination with several other factors, additively contributes towards RAG transcriptional activation, but is not essential. Second, BSAP+/− mice may compensate for the loss of BSAP expression through activity of other, either known or unknown Pax family members. Finally, it is possible that BSAP is necessary for regulated RAG expression only during later stages of development. Since the second to fourth RAG expression waves of RAG (in pre-, immature, and mature GC B cells, respectively) come after the BSAP+/− B developmental arrest, this phenotype cannot rule out a role for BSAP later in B cell development, a point that will be further discussed below.

Although the effect of BSAP deficiency on RAG expression in later B cell differentiation cannot be determined in the BSAP+/− phenotype (which could be tested by rescuing B cell differentiation with Ig transgenes), the presence of RAG transcripts and accompanying HC rearrangements suggests that
BSAP is not required for RAG activation in pro-B cell development [369]. Additionally, another gene that can also be re-expressed along with RAGs in GCs is λ5. This gene is also a potential BSAP target, based on the existence of BSAP-binding sites in its promoter [248]. Further indirect evidence for a role for BSAP in RAG reactivation in later B cell development comes from studies of IL-7 signaling in the periphery. As mentioned in the General Introduction, IL-7 was recently shown as a co-factor of RAG activation at this stage in vivo and in vitro [199]. In this context, IL-7R⁺ mice have reduced BSAP transcript levels and have a B cell developmental block at the same stage as BSAP⁻ mice [437]. This would suggest that BSAP functions downstream of an IL-7R signaling pathway. Indirect evidence also comes from the important roles of BSAP established in various IL-4+LPS-mediated processes which occur in parallel with RAG re-expression in the periphery. These include studies where overexpression of BSAP in plasma cells or resting splenic B cells induce proliferation, whereas BSAP downregulation by specific antisense oligonucleotides inhibit LPS+IL-4-mediated proliferation and class switching [438]. Similar approaches have also shown that J chain and α or ε HC transcription are dependent on BSAP expression [373, 439-441]. Several groups have also shown that a conserved BSAP site in the mouse and human Ige promoters is required for their LPS+IL-4 inducibility [442-444].

Transcription factors upstream of BSAP may also be directly or indirectly implicated in RAG regulation. One particularly interesting candidate is the E12/E47 protein complex, whose properties were discussed in the General Introduction. E2A⁺ mice have a block in B cell development in Hardy fraction A cells, an arrest which is not only earlier than in RAG1⁺ or RAG2⁺ mice, but also than in Pax-5⁺ mice (Figure 1.7) [207, 208]. Furthermore, E2A⁺ mice besides lacking transcripts for RAG1 and the B-cell specific genes λ5, CD19, and mb-1, also lack those for BSAP [208]. Therefore, the E2A proteins may either have a direct role in transcriptional activation of the CD19, λ5, and RAG loci and/or may work indirectly by activating BSAP, which in turn activates transcription of RAG and other B cell-specific genes. Other B cell-specific transcription factor candidates upstream of BSAP may also be involved in RAG activation, based on their expression pattern and knockout phenotypes. These include EBF, Sox-4, and LEF-1. For example, EBF-deficient mice have a block early in B cell development.
development similar to E2A<sup>+</sup> mice, which is prior to that of not only RAG<sup>+</sup> mice, but also the BSAP<sup>+</sup> phenotype.

Yet other B-cell specific transcription factor candidates upstream of BSAP may be involved in the B-cell specific repression of RAGs. As mentioned in the General Introduction, Id and ZEB may be indirect candidates for RAG transcriptional repressors as direct E2A regulators. However, the pro-B restricted expression pattern of Id mRNAs suggests that Id-mediated E2A/RAG repression occurs only as a pro-B cell developmentally-restricted mechanism (see Table 1.1; [23]). Since BSAP may or may not function in earlier B cell-specific RAG activation, it therefore remains to be determined if Id is linked to BSAP regulation as an upstream target. This indirect negative regulatory effect on RAG transcription further demonstrates that complex transcriptional networks, like those involved in Ig transcription, are involved in regulating the RAG genes. Furthermore, since expression of the RAG and Ig genes are both lymphoid and developmentally-restricted, it is not surprising that both of these loci share common transcription factors such as E2A and Id.

5.3.2 Emergence of an increasingly complex picture of RAG regulation: further evidence for distinct regulation of RAG expression in T and B lineages

The fact that there are potential RAG regulatory factors (such as hBRAG and BSAP) expressed preferentially in B cells is suggestive in itself that there is differential regulation in T and B cell lineages. In addition, although T cells may also undergo receptor editing in DP TCR<sup>+</sup> thymocytes [224, 445] and RAG is re-expressed in peripheral CD4<sup>+</sup> T cells [74], the signals for RAG re-induction in B cells, such as IL-4+LPS induction and CD40 ligation are unique to the B lineage. Additionally, the fourth wave of RAG re-expression is also regulated differentially in the context of respective antigen receptors; in the presence of BCR crosslinking, RAG expression is shut off in GC mature B cells [253], whereas peripheral T cells appear to reactivate the V(D)J recombinase in the analogous setting [74].

Whatever mechanisms involved, there are several further lines of evidence supporting differential RAG regulation in T and B cells. Lauring and Schlissel have recently determined that the mouse RAG2
promoter can be regulated distinctly in B and T cells [271]. In particular, the putative BSAP-containing cis region found critical for promoter activity in B cells has an overlapping portion determined to be important for T cell-specific activity via interaction with an unknown factor, distinct from BSAP. Furthermore, for full T cell-specific RAG2 promoter activity, a region further downstream in the RAG2 promoter was also required. Upon deletion of this region, T, but not B cell activity is affected. The B lineage-specific transactivation of the RAG2 promoter by BSAP is consistent not only with its expression pattern, but also the phenotype of its targeted disruption, which affects B but not T cell development. Further evidence for distinct RAG regulation mechanisms in B and T lineages comes from recent in vivo developmental rescue and transgenic reporter experiments [276, 277]. In the rescue experiments, several genomic constructs (containing sequences 5' of the RAG2 TSS) that are competent for rescuing RAG2 expression in pre-B cells have been subsequently transfected into RAG2+ ES cells and tested for rescue of lymphocyte development in RAG2 mouse chimeras [276]. It was found that relative to constructs containing larger portions 5' of the RAG2 promoter, more proximal elements (2-7 kb 5') rescue B, but not T cell development, suggesting the presence of distal T cell-specific cis regulatory elements. In the transgenic reporter experiments, numerous BAC transgenes spanning regions flanking the RAG locus were generated that contained distinct GFP and YFP fluorescent indicators in place of the RAG1 and RAG2 genes, respectively [277]. Using this approach, it was found that coordinate expression of RAG1 and RAG2 in T versus B cells is regulated by distinct cis elements on the 5' side of RAG2. Overall, the above reports suggest that lineage-specific regulation of RAG1/2 is not only controlled by a separate set of T and B transcriptional regulatory factors and/or signaling pathways within the RAG promoters, but also by different cis-acting regulatory regions outside of the RAG promoters.

Additional evidence for distinct B and T cell RAG regulatory mechanisms comes from E2A knockout and transgenic mouse data. Specifically, while the E2A knockout experiment indicates that the E12/E47 products are crucial for RAG activation (directly or indirectly) in early B cells, it also shows that the E2A products are not crucial for RAG expression and TCR rearrangements, since E2A- mice continue to express RAG and TCRα/β transcripts in thymocytes [207, 208]. Furthermore,
introduction of E2A into a pre-T cell line stimulates the expression of RAG1/2. Therefore, RAG1 activation in thymocytes is not dependent on E2A whereas its expression in developing B-lymphocytes is. Additionally, this also indirectly links the Id protein in the B cell specific negative regulation of RAGs since, as mentioned above, this protein appears to negatively-regulate the E2A proteins.

5.3.3 An overall model of hBRAG, BSAP, and other potential regulatory factors involved in early and late B cell-specific RAG regulation

Based on the analysis of E2A, BSAP, Id, and hBRAG function, combined with what is known about the signaling-mediated regulation of RAGs (outlined in section 1.5.2.2 of the general introduction) and on the speculations in the above two sections, Figure 5.3 summarizes an overall model for the hierarchical relationship of the various factors involved in early and late B cell-specific regulation of the RAGs. In this model, the E2A transcription factor is a key B cell-specific regulator of RAG throughout development in this lineage.

In early B cells (pro- and pre-B cells) specifically, signals through CD19 and the pre-BCR may have a role in altering the IL-7 driven induction of RAG expression. The relationship between these surface signals and other downstream factors is not known, but they may be involved in LEF-1, Sox-4, and EBF transcription factor activation, themselves potential RAG activators which are specifically expressed at this stage. The Abelson Tyrosine Kinase may also have a role in regulating early B cell-specific RAG-inducing transcription factors, including NF-κB. Conversely, Id may play a role in RAG repression at this stage.

In mature B cells, BCR ligation can deliver negative or positive signals, and further downstream, BSAP may play a specific role in RAG induction. Since BSAP is an E2A and an EBF target (based on the E2A+ and EBF+ phenotypes) it is presumably part of an EBF/E2A RAG activation pathway, although these factors may also act independently of each other and of BSAP. Furthermore, the combination of IL-4 and CD40 can re-induce RAG expression in GCs. It is not known if and how differential signaling through BCR, or through cytokines/co-stimulatory receptors affect E2A and BSAP transcription factor activity. IL-7R signaling may also have a role in GC RAG re-expression in a BSAP-dependent fashion. As has been discussed above section for BSAP, it can be hypothesized
Early B cell-specific factors

Late B cell-specific factors

Figure 5.1. Speculative model for hierarchy of factors involved in the early and late B cell-specific regulation of RAG transcription. In this model, hBRAG is shown in both early and late development either as a cell surface receptor delivering signals to any one of the various downstream factors shown (including BSAP and/or E2A) or via other as yet unidentified factors/pathways. Alternatively, hBRAG is shown as an intracellular enzyme involved in transport/synthesis/post-translational modification of such factors. In early B cell development, RAG regulation may include contact via the candidate stromal cell receptor C2.3 receptor candidate, and analogously, hBRAG could be a pro- or pre-B cell-specific receptor for a stromal cell signal. In mature B cells, in contrast to early B cell development, Id is not involved in RAG regulation. Furthermore, instead of stromal cell contact, the combination of various T cell signals such as gp39 (CD40L), IL-7, and IL-4 are involved in RAG induction and APCs like FDCs efficiently present antigen so that high avidity cross-linking can either shut off RAG expression in GC B cells, or in immature BM B cells, the opposite effect is seen (RAG induction). As with early B cell signaling, it is not clear how these signaling-mediating factors would affect activity of downstream transcription factors. Note that only B-cell specific factors are shown, although other RAG-regulating factors/pathways common to both T and B cells (which may or may not be interconnected with B-cell specific factors; for example PKC/Ca^{2+}, cAMP signaling pathways, phosphatases 1A and 1B, etc.) are not shown for sake of simplicity (see Figure 1.4 and Tables 1.3 and 1.4 of General Introduction for complete list of negatively and positively-regulating RAG factors). Arrows indicate the transcriptional targets of the E2A and the BSAP products, which in early B cell development, may also control the transcription of other early B cell specific genes including VpreB, λ5, CD19, and the IgH and κ chain loci. The Id proteins which function to inactivate bHLH proteins such as the E2A products, have been included in the diagram as indirect repressors of RAG transcription in pro-B cell development. Question marks represent uncertainty as to whether E2A proteins function directly or indirectly as RAG transcriptional activators.
that hBRAG also has a preferential role in RAG regulation in later B cell development (including the induction of the fourth wave of RAG reexpression), based on several clues. First, the identification of hBRAG as a RAG co-expressed gene via differential display of clonally related GC variants, like BSAP, has not only provided evidence that it may regulate the endogenous RAGs [392], but also suggests its role in the context of GC receptor reexpression/receptor revision, since the mature cell line used in this study undergoes spontaneous secondary LC rearrangements in culture [223] and can inducibly upregulate RAG expression via BCR cross-linking or second messenger stimulation [75]. Additionally, gene transfer experiments of hBRAG in the mature GC B cell line OCI LY8-C3P increases RAG1 mRNA levels [49]. Thirdly, this gene's expression is particularly high in fetal and adult secondary lymphoid tissues enriched in mature B cells (spleen, lymph nodes, and PBL; [49]), sites where receptor revision/RAG re-expression have been demonstrated. Furthermore, type II molecules with immune function, as described earlier, are traditionally involved in signaling in later B cell differentiation. This is also consistent with the evidence I have provided in Chapter 4 supporting a function for hBRAG in BCR-associated signaling.

An interesting question that emerges based on the above model, is if the novel cDNAs I have identified by DD RT-PCR, including hBRAG (as either a B cell-specific surface receptor and/or as an enzyme involved in synthesis/post-translational modification of other B cell-specific factors) has a potential relationship to BSAP and/or E2A in the B cell-specific regulation of RAGs. In this context, it would be of particular interest to determine if BSAP and E2A are targeted by hBRAG or whether this presumably upstream factor acts via a separate pathway. As a preliminary experiment, the Northern blots of hBRAG stable transfectants in both K562 and OCI LY8-C3P (see Chapter 3, Figure 3.9) could be re-probed with E2A and BSAP-specific probes. This could be further tested for by transient transfection of hBRAG into pro-B cells from Pax-5 or E2A- mice. This experiment, however, would assume that the control transient transfection (hBRAG into pro-B cells from WT mice) results in a similar increase in RAG1 mRNAs as that seen for stable transfection into OCI LY8-C3P. Since I and others have shown that RAGs can also be post-transcriptionally regulated, prior to doing the aforementioned experiment, it may be worthwhile to see at what levels the hBRAG-mediated increases
are occurring. This could be done by nuclear run-on analysis and/or RNA stability assays, similar to those I performed to assess the regulatory mechanism by which BCR ligation mediated the RAG increases in OCI LY8-C3P [75].

The concept of RAG expression not only being regulated by lineage-specific factors and elements, but also by distinct stage-specific factors/elements within these lineages (such as those shown in Figure 5.3), may help to resolve an intriguing contradiction. In particular, one recent report using the BAC transgenic reporter system (described in the previous section) has found that there is no re-induction of RAG expression after immunization [446], in contrast to numerous previous reports [71, 249, 253]. One interpretation for this discrepancy is that while the cis-region within this transgene that is sufficient to drive appropriate, constitutive RAG expression in various B cell developmental stages does not contain factors and/or elements required for RAG re-induction in late B cell differentiation.

5.3.4 Implications for altered hBRAG and BSAP gene expression in potential B-cell-specific immunodeficiencies, cancers, and autoimmunities.

Because the expression of RAG1/2 is tightly controlled, it is not surprising that deregulation of these genes (either under- or overexpression) has deleterious effects to the lymphocytic system of mouse and man. In particular, as will be discussed below under- and overexpression can result in inefficient antigen receptor gene recombination and deleterious recombination errors, respectively. This emphasizes the critical importance of precise RAG regulation in B and T lymphocyte development and makes an understanding of this precise regulation pivotal in the prevention/treatment of transformations and immunodeficiencies. Further B and T lineage-specific immunodeficiencies may be due to RAG control factors which are B or T lineage-specific. This aspect will be discussed below in the context of B cell-specific defects in RAG expression that potentially stem from deregulated hBRAG and BSAP expression.

Potential diseases caused by underexpression of RAGs, BSAP and hBRAG

As mentioned in the General Introduction, RAG1 or RAG2<sup>+</sup> mice are viable but have a SCID phenotype: small lymphoid organs that fail to produce mature B and T lymphocytes (Figure 1.6; [50,
In neither mouse are other developmental abnormalities found. The disfunctioning and/or underexpression of human RAG products results in an analogous phenotype to that of the RAG* mouse. In particular, missense mutations in either the RAG1 or RAG2 proteins have recently been linked to two SCID states: Ommen Syndrome (OS) and B-SCID [363, 364]. The mutations reported to date, which are estimated to account for about 14% of total human SCID cases, specifically affect the structure of the proteins, not their expression levels or subcellular localization. These include truncations due to premature stop codons and alterations in critical core domain residues leading to abrogated associations with each other and with the RSS, respectively. As expected, the defective RAG proteins from patients all had a decreased ability to rearrange artificial extrachromosomal substrates and cleavage activity in vitro.

The proportion of human SCID cases due to RAG defects may be an underestimate, as other B-SCID could be due to mutations in the tissue-specific regulatory regions and protein factors affecting RAG mRNA and/or protein expression in pathways common to B and T lymphocytes. In addition, mutations in B and T lineage-specific RAG regulatory factors would not manifest themselves as SCID phenotypes like those seen in RAG* mice or OS/B-SCID patients but rather as B or T lineage-specific immunodeficiencies, and thus would not be accounted for by the above studies. In this context, mutations causing underexpression of factors such as hBRAG and BSAP, which may be involved in a pathway positively regulating RAG mRNA expression in B-cell specific context could result in B-cell specific immunodeficiencies. In particular, while BSAP deficiency does not seem to affect RAG mRNA expression (at least in BSAP* mice; [369]), deficiency in hBRAG, as a potential RAG regulator in antigen-independent B cell differentiation, could potentially have a phenotype similar to the XLA defect. This defect, caused by mutation in the Btk gene, results in a specific and nearly complete block at the pre-B cell stage [34, 35]. A potential involvement of both hBRAG or BSAP in RAG regulation later in B cell development could lead to peripheral B-cell specific immunodeficiency/function because of defective GC receptor revision. With respect to hBRAG as a signaling receptor, defective could also be due to defects in more general B cell-specific signaling functions. For example, if hBRAG is involved in mature B cell co-stimulation, a defect in its gene
may phenotypically manifest itself similarly to those seen in Hyper IgM (HIM) patients, who have a defect in the typeII molecule CD40L. The developmental block occurs in peripheral B cell pools, which are rendered largely non-functional due to defective GC formation, isotype switching, affinity maturation, and plasma and memory cell differentiation [27-29]. As a result, HIM patients have a severely compromised humoral immune system, in particular, high levels of IgM but markedly reduced levels of other Ig isotypes. CD40L− mice also show a similar phenotype [80, 447].

**The possible involvement of BSAP and hBRAG in potential disease states resulting from RAG overexpression**

In the mouse, Wayne et al. made transgenic mice in which RAG1/2 were overexpressed via placement under the control of the lck promoter [58, 59]. Deregulated expression of either gene alone had no significant effect but their co-ordinate deregulation resulted in late T cell developmental abnormalities. These mice developed enlarged spleens and lymph nodes and had abnormal accumulations of lymphocytes in perivascular locations of most organs. These alterations were associated with impaired cellular and humoral immune responses. In the human, the overexpression of the RAG products have not yet been directly linked to any disease states. However, since V(D)J recombinase activity has been correlated to chromosomal translocation events involving the misjoining of Ig/TCR gene segments with proto-oncogenes in B/T lymphomas, respectively, there is strong circumstantial evidence implicating the involvement of RAG expression in these events.

In this context, two mechanisms for how RAG-mediated translocation events may occur have been proposed, based on the sequence analysis of various translocation breakpoints. First, an interchromosomal form of "normal" V(D)J recombination can occur such that one breakpoint region is at an antigen receptor locus, at or near an authentic RSS site (typically at a D or J segment of the IgH or TCR α, β, or δ loci) which can be recombined to a heptamer or nonamer-like RSS sequence in a proto-oncogene (such as bcl-2 or c-myc) on a partner chromosome [287, 448-452]. Because there are cases where only single RSS appear to be involved (for example in t(14:18) B cell lymphomas, where RSS-like sequences cannot be identified at the breakpoint on the partner chromosome), these cannot be explained by standard V(D)J recombination events [453-455]. In light of these observations coupled
with recent evidence that signal ends can be diverted into a RAG-mediated transpositional reaction [456, 457], a second mechanism hypothesized by Hiom and Gellert is that inappropriate transposition events underlying such aberrant chromosomal rearrangements and resulting to lymphoid neoplasms could be promoted by the Rags [456]. The RAGs as transposases and their potential roles in chromosomal translocations are beyond the scope of this discussion and will therefore not be discussed further here; the reader is consulted to read some recent reviews on the subject [458-460].

It can be hypothesized that the frequency of RAG-mediated interchromosomal V(D)J recombination- and transposition-like translocation events is small in cells expressing RAG "normally". In such cases, advantages gained by a diversified lymphocyte repertoire greatly outweigh the risk of generating chromosomal translocations and hence a malignant phenotype. However, in cells with deregulated and/or overabundant RAG1 and/or RAG2 expression (including outside the G1/G0 cell cycle boundary and/or at stages when the RAG "waves" are shut off), increased translocation events leading to lymphomagenesis may occur due to alterations in the two aforementioned mechanisms. First, there would be an increased presence of potentially dangerous reaction intermediates (signal ends) used as substrates for chromosomal translocation events including secondary rearrangement-induced translocation events in other mature B cell subsets. Secondly, it may cause a shift in the relative abundance of V(D)J components from a pathway where the relative expression of DSBR components predominate (thereby allowing proper resolution/end-joining of coding and signal ends, and suppressing transposition pathways/events) to a pathway where these components are underrepresented and in which the RAGs are subsequently transpositionally active [461]. Hence, RAG expression itself may be one of the factors affecting the balance of the V(D)J recombination, DNA repair, and transposition pathways in the cellular environment.

Because hBRAG can upregulate RAG mRNAs in the context of mature B cells, and BSAP has been shown to confer B cell-specific regulation of the RAG2 promoter, it follows that mutations leading to overexpression of either factor may correlate with and lead to increased incidences of putative RAG-mediated B cell leukemias. Based on the data presented in this thesis, hBRAG is expressed in all B lymphopoietic tissues, but at especially high levels in peripheral tissues, and may therefore be
important in the progression of malignancies at later stages of B cell development. Furthermore, high mRNA levels of BSAP and hBRAG can be co-expressed with those of RAG in pro-B, pre-B, and mature B lines, and may therefore have the potential to be deregulated and to cause translocations at any of these stages (See figure 1.3 to see where the various malignancies occur relative to normal stages of human B cell differentiation). In this context, many of the BSAP*hBRAG*RAG* pre-B and mature B cell lines assessed in these studies are pre-B ALL, follicular, and diffuse large cell lymphomas, and as such, bear evidence of the translocation events described above. Assuming cell surface expression, hBRAG in particular could be used as a potential clinical marker of B leukemia etiology, including B cell lymphomas resulting from interchromosomal V(D)J recombination and/or transposition-like translocation events. Furthermore, while hBRAG and BSAP are normally not expressed in the T cell lineage, they may be under "deregulated" conditions, and could therefore also potentially have a role in the etiology of various T cell lymphomas. The molecular cloning of hBRAG product will allow it to be tested in this regard by its enforced overexpression in transgenic mice, for example.

Overexpression or underexpression of factors involved in the regulation of RAGs may also affect the balance between immune diversity and autoimmunity. In particular, inappropriate "turning on" of RAGs, mediated by irregular expression of regulatory factors including hBRAG or BSAP (or E2A, Id, etc...) could hypothetically increase the amount of secondary rearrangement events. In the context of GC receptor selection, this may in turn result in the diversification of receptors normally incapable of being modified, thereby increasing the potential for autoreactive specificities. Conversely, shutting off RAG expression via irregular expression of such RAG-regulatory factors may reduce secondary rearrangement events. With respect to receptor editing and GC receptor selection, this would presumably result in the inability to rescue certain bone marrow Ig receptors or to diversify GC Ig receptors, respectively. In either case, this may result in more restricted Ig diversity.
5.4 Concluding remarks

In using the screening methodology described in this thesis for identifying RAG1 co-expressed genes, I postulated that a fraction of the cDNAs obtained would correspond to lineage and/or stage-specific RAG-regulating genes (either known or novel). Out of the three DD RT-PCR amplicons obtained in these screenings with corresponding known, full-length sequences (RAG2, hBRAG, and BSAP), two (hBRAG and BSAP) are not only likely involved in the regulation of RAG1, but are also preferentially expressed in B cells. These overall results suggest that the corresponding genes of some of the other uncharacterized novel amplicon sequences obtained are also somehow involved in B cell-specific RAG regulation. Furthermore, the possibility of several genes with related function (as RAG regulators) is consistent with the prediction that differential gene expression in OCI LY8 variants is specific to control factors mediating constitutive differences in RAG1/2 expression and the absence of other observable phenotypic differences [264]. Furthermore, the previous finding that the constitutive differential expression in these variants is de novo synthesis-dependent [264], suggests that the two amplicons that were found to inversely correlate with RAG1 expression (2A1 and 4A1) are good candidates as corresponding to a putative differentially-expressed negative RAG regulator.

Characterization of the novel amplicons will first have to include the acquisition of corresponding full-length cDNAs, either by 5' RACE, (for amplicons detecting transcripts which are ≤1 kb such as 5C1, 5C2, and 5C3), or by cDNA library screening (for amplicons detecting larger transcripts). Larger IMAGE consortium cDNA clones corresponding to the 4A1 and 5C2/5C3-1.0 amplicons can also be used, either directly to define the potential ORFs of corresponding complete cDNAs, or used as probes in library screening and Northern blotting. Assuming these as-yet uncharacterized clones have roles in the processes of RAG regulation and B cell development, it will be interesting to assess if they function in the same pathway that hBRAG and/or BSAP are involved in. If so, this would further reinforce the RAG regulation-specific differences between the OCI LY8 clonal variants, since it would be highly improbable to find differentially expressed genes with not only related functions, but that are also part of a common pathway had large numbers of non-specific differences been present.
REFERENCES
References

References


References


74 McMahan, C. J. and Fink, P. J., RAG reexpression and DNA recombination at T cell receptor loci in peripheral CD4+ T cells. *Immunity* 1998. **9**: 637-647.


References


101 Takemori, T., Mizuguchi, J., Miyazoe, I., Nakanashi, M., Shigemoto, K., Kimoto, H., Shirasawa, T., Maruyama, N. and Tanaguchi, M., Two types of μ chain complexes are expressed during differentiation from pre-B to mature B cells. EMBO J. 1990. 9: 2493-2500.


105 Loffert, D., Ehlich, A., Muller, W. and Rajewsky, K., Surrogate light chain expression is required to establish immunoglobulin heavy chain allelic exclusion during early B cell development. Immunity 1996. 4: 133-144.


111 Horne, M. C., Roth, P. E. and DeFranco, A. L., Assembly of the truncated immunoglobulin heavy chain Dmu into antigen receptor-like complexes in pre-B cells but not in B cells. Immunity 1996. 4: 145-158.


242
References

References


References


245


References


Menetski, J. and Gellert, M., V(D)J recombinase activity in lymphoid cell lines is increased by agents that elevate cAMP. *Proc. Natl Acad. Sci. USA* 1990. 87: 9324-9328.


References


264  Verkoczy, L. K., *Immunology*, University of Toronto, Toronto 1995, p. 95.


250
References

References


252
References


253


934 Kozak, M., Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 1986. 44: 283-291.
References


425 Miele, G., MacRae, L., McBride, D., Manson, J. and Clinton, M., Elimination of false positives generated through PCR re-amplification of differential display cDNA [In Process Citation]. Biotechniques 1998. 25: 138-44.


258
References


450 Tsujimoto, Y., Louie, E., Bashir, M. M. and Croce, C. M., The reciprocal partners of both the t(14:18) and the t(11:14) translocations involved in B-cell neoplasms are rearranged by the same mechanism. Oncogene 1988. 2: 347-351.


APPENDIX

Appendix
Figure A.1. Comparison of KIAA0598 and hBRAG cDNA sequences. A. Schematic representation of KIAA0598 and hBRAG complete cDNAs. Single base pair additions in either cDNA (relative to the other) are shown as black vertical lines. The open reading frames are shaded gray. B. Alignment of KIAA0598 and hBRAG 5' UTR and coding region portions of cDNA sequences.
Figure A.2. Alignment of the hBRAG and KIAA0598 protein sequences. Complete identities are shown by stars and partial identities (similarities) are shown by dots. The accession number for the KIA0558 protein is AB011170.
Figure A.3. Immunochemical characterization of rabbit polyclonal antiserum raised against N and C-terminal hBRAG peptides. A. Reactivity of rabbit polyclonal antiserum to N and C-terminal peptides. Shown are the 7th week and 13th week bleeds (1st and terminal bleeds, respectively) of crude antisera raised against the hBRAG N-terminus (lots 2421, 2422) and C-terminus (lots 2423, 2424), or of various corresponding affinity-purified antibodies. The relative reactivity to BSA-peptide conjugate strips is representatively shown by using 1:10000 dilutions of all antibodies, including pre-immune negative controls towards 10 µg of the corresponding N and C-terminal BSA conjugated peptides UT952 and UT953, respectively. B. Detection of separate specific immunoreactive bands by single and dual peptide competition Western blot assays of membrane strips prepared from endogenous hBRAG A8-6P cell line lysates using α-hBRAG antibodies, affinity purified from 2422 and 2424 antisera raised against N and C-terminal hBRAG peptides, respectively. Immunoreactive bands specifically detected by 2422 and 2424 antibodies are denoted by * and o, respectively. In all instances, 1 µg/ml of affinity purified antibodies was used. Lanes are as follows: lanes 1-5; incubation with affinity purified N-terminal specific α-hBRAG antibodies alone or in the presence of 10 ng, 100 ng, 1 µg, or 10 µg/ml of N-terminal competing peptide UT952, lanes 6-10; incubation with affinity purified C-terminal specific α-hBRAG antibodies alone or in the presence of 10 ng, 100 ng, 1 µg, or 10 µg/ml of C-terminal competing peptide UT953, lane 11-20; incubation with a mixture of N and C-terminal specific α-hBRAG antibodies alone (lane 11) or with preincubation with N-terminal competing peptide UT952 at 100 ng, 1 µg, and 10 µg/ml (lanes 12-14), preincubation with C-terminal competing peptide UT952 at 100 ng, 1 µg, and 10 µg/ml (lanes 15-17), or preincubation with both N- and C-terminal competing peptides UT952 and UT953 at 100 ng, 1 µg, and 10 µg/ml (lanes 18-20). For panels B-D, membrane strips in all cases contained equal protein loads (40 µg) and lysates were run under denaturing, non-reducing SDS-PAGE conditions.
### A

<table>
<thead>
<tr>
<th></th>
<th>7th week bleed</th>
<th>12th week bleed</th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-immune sera</td>
<td>#2421 #2422 #2423</td>
<td>#2421 #2422 #2423</td>
</tr>
<tr>
<td>crude sera</td>
<td>#2422 #2423</td>
<td>#2422 #2423</td>
</tr>
<tr>
<td>flowthrough</td>
<td>#2421 #2422</td>
<td>#2421 #2422</td>
</tr>
<tr>
<td>affinity purified fraction</td>
<td>#2424</td>
<td>#2424</td>
</tr>
</tbody>
</table>

### B

<table>
<thead>
<tr>
<th>α-hBRAG antibody blocking peptide</th>
<th>N-terminal, 2422</th>
<th>C-terminal, 2424</th>
</tr>
</thead>
<tbody>
<tr>
<td>UT952</td>
<td>UT953</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N-terminal, 2422 + C-terminal, 2424</th>
</tr>
</thead>
<tbody>
<tr>
<td>UT952</td>
<td>UT953</td>
</tr>
<tr>
<td>UT952 + UT953</td>
<td>UT953</td>
</tr>
</tbody>
</table>

![Image of gel electrophoresis results](image-url)
Figure A.4. Reactivity of N and C-terminal hBRAG antibodies to in vitro-translated hBRAG cDNA products.

N-glycosylation of the in vitro-translated hBRAG cDNA and recognition of glycosylated and unglycosylated hBRAG isoforms by N-terminal, but not C-terminal α-hBRAG antibodies. Autoradiogram of the [35S] methionine-labeled cell-free translated hBRAG gene products either directly resolved by SDS-PAGE analysis (lanes 1-5), or immunoprecipitated with either affinity purified α-hBRAG antibodies from terminal bleed antiserum raised against the N-terminal hBRAG peptide UT952 (2422; lanes 6-9), or the C-terminal hBRAG peptide UT953 (2424; lanes 10-14) prior to SDS-PAGE. The open arrow shows the unglycosylated hBRAG doublet product and the closed arrow shows the ~70 kDa glycosylated product shown by a closed arrow (see text for details of relative amounts of glycosylated and unglycosylated products in whole and in immunoprecipitated cell-free fractions). Numbers to the left indicate the migration positions of the protein standards. Lane 1 represents a negative control in which no cDNA was added. Translation reactions were carried out either without microsomes (lanes 2, 6, 11), with microsomes only (lanes 3, 7, 12), with microsomes and the competitive inhibitor NYT (lanes 4, 8, 13), or with microsomes only and digestion with PNGaseF after translation (lanes 5, 9, 14). The designation “S” indicates the linearized pBluescript 8-3 hBRAG that was transcribed in vitro, using the T3 promoter, to generate a sense RNA, whereas “AS” indicates in vitro transcription of 8-3 hBRAG with the T7 promoter to generate an anti-sense RNA.
<table>
<thead>
<tr>
<th></th>
<th>whole</th>
<th>IPs: 2422</th>
<th>IPs: 2424</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>-</td>
<td>AS</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8-3 cDNA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>microsomes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PNGase F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYT</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

kDa

194 120 87 64 52 39 26 21

1 2 3 4 5 6 7 8 9 10 11 12 13 14