Functional study of the cis-elements in the IgH locus using Bacterial Artificial Chromosomes

Buyi Zhang

A dissertation submitted to the Graduate Faculty in Biochemistry in partial fulfillment of the requirements for the degree of doctor of philosophy, The City University of New York

2005

UMI Number: 3187357

UMI Microform 3187357

Copyright 2005 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

This manuscript has been read and accepted for the Graduate Faculty in Biochemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

Date	Chair of Examining Committee Dr. Laurel Eckhardt, Hunter College
Date	Executive Officer Dr. Lesley Davenport
	Dr. Benjamin Ortiz, Hunter College
	Dr. David Foster, Hunter College
	Dr. Barbara Birshtein, Albert Einstein College of Medicine
	Dr. Andrea Cerutti, Weill Medical College of Cornell University

Supervisory Committee

The City University of New York

Abstract

Functional study of the cis-elements in the IgH locus using

Bacterial Artificial Chromosomes

By

Buyi Zhang

Adviser: Professor Laurel Eckhardt

Cis-elements play important roles in recombination, transcription, and hypermutation events that take place in the IgH locus. The intronic enhancer, Eμ is indispensable for V to DJ recombination. The 3'IgH enhancer pair hs3b and hs4 is required for class switch recombination (CSR). Copious transcription from the IgH locus in Ig secreting cells is dependent on an intact 3'IgH region. In this study, reporter genes with the entire 3'IgH regulatory region were constructed using bacterial artificial chromosomes (BACs). These reporter constructs have the advantages of having all the cis-elements in their natural positions and orientations and preserving the possible higher order DNA structure. We found that the 3'IgH regulatory region could confer transgene transcription in most integration sites in an Ig-secreting cell line, but the expression level was not dependent on the copy number of the transgene. However, Eμ together with the 3'IgH regulatory region could enable the transgene to be expressed in a copy number dependent manner, suggesting that all the elements in this particular BAC constitute an LCR. In addition, we showed that hs4 deletion from the 3'Igh regulatory region did not affect the ability of this region to confer transgene high level transcription in the same integration sites, suggesting that there is more redundancy in function than previously appreciated in the IgH locus and that hs4 is dispensable for maintaining IgH gene transcription in Ig-secreting cells, even in the absence of Eμ.

Acknowledgments

First, I would like to thank my mentor, Dr. Laurel Eckhardt. She has high expectations and demands excellence from her students. She is particularly sharp in solving complicated problems and gaining insights from data. I was lucky to have the opportunity to work with her during these years. Under her guidance, I learned to prioritize tasks and always focus on the most important ones and to keep a good balance between quality and productivity.

I would also like to thank Dr. Ortiz for his expertise in LCR and his advice on experimental design and Dr. Foster for very valuable and practical advice on productivity which not only has contributed to the completion of my Ph.D study but also will be helpful for my success in the future.

Current and past members of the Eckhardt's lab have lent me tremendous help on the completion of the projects. Special thanks to Adrienne Petrillo, Fubin Li, Yi Yan, , Maria Kon, Pavel Munerman, Steven Williams, and David Miller.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Abbreviations:

3' IgH RR: 3' IgH Regulatory Region

BAC: Bacterial Artificial Chromosome

CSR: Class Switch Recombination

EMSA: Electrophoretic Mobility Shift Assays

GR: Glucocorticoid Receptor

IgH: Immunoglobulin Heavy Chain

IP: I Exon Promoter

HSC: Hematopoietic Stem Cell

HSs: DNaseI Hypersensitive sites

LCR: Locus Control Region

MARs: Matrix Attachment Regions

PEV: Position Effect Variegation

SHM: Somatic Hypermutation

Chapter 1 General Introduction

Our lab uses the IgH locus as a model system to study the mechanisms that activate and sustain transcription of tissue specific genes. Thus, it is necessary to introduce the organization of the IgH locus, stages of B cell development and the various events that occur at the IgH locus during B cell development. Then the issue of the mechanism of transcription activation at the IgH locus which concerns the core of the thesis will be discussed. Transcription activation has become much more complex than what is shown in the simplest model as on and off switch. Although enhancer activities in various cis-elements in the non-coding region of the genes play an important role in transcription regulation, other mechanisms, such as chromatin structure, have been clearly shown to be critical in regulating the choice of the gene to be transcribed as well as the rate of transcription. Recently, a lot of effort has been devoted to defining locus control regions (LCR) in tissue- and stage- specific genes. Defining an LCR in the IgH locus is important to completing our knowledge of transcription regulation, in aiding our quest into the mechanisms of recombination and hypermutation events, and in providing a framework for dissecting the functions of individual cis-elements, or even specific protein binding motifs within these cis-elements. Therefore, chromatin modification and the concept of LCR will be introduced.

1.1 IgH gene structure and B cell development

1.1.1 The organization of the murine IgH locus

Antibodies are the major players in the humoral immune response. They protect the host organism by either neutralizing toxic antigen, binding to and thereby tagging

microorganisms for destruction by phagocytic cells, or recruiting the complement system to lyse foreign microorganisms. An antibody molecule consists of four polypeptides, two identical short ones called light chains and two identical long ones called heavy chain. The structure of an antibody is shown in Figure 1-1,

The assembly and expression of the IgH heavy chain gene of the antibodies is much more complex than the expression of most of the genes in living cells in living organisms. The germline IgH locus spans about 3Mb and is located close to the telomere of chromosome 12 (mouse) and chromosome 15 (human), respectively. There are four clusters of donor sequences for the coding sequences of the heavy chain gene. They are (from upstream—telomere to downstream-centromere) in the order of the V_H cluster, D_H cluster, J_H cluster, and the cluster of constant regions (Figure 1-2).

In the murine IgH locus, there are 14 V_H families based on sequence similarity (Kofler et al, 1992), with the J558 family—the largest V_H family—residing at the 5' (close to telomere) and the 7183 family residing at the most 3' of the V_H cluster (Brodeur, 1988). Upstream of each V gene, there is a leader sequence preceded by a lymphoid specific V_H promoter. Downstream of the V_H cluster is the D cluster consisting of three families of D_H genes, and the J cluster consisting of J_H genes (Honjo and Matsuda 1992).

Figure 1-1 Schematic structure of an antibody An antibody molecule consists of two identical heavy chains and two identical light chains. Each heavy and light chain in an antibody molecule contains an amino-terminal terminal variable (V) region which determines the specificity of the antibody and constant (C) region which determine the effector function of the antibody. Both heavy and light chains are held together by disulfide bonds.

Figure 1-1 Schematic structure of an antibody

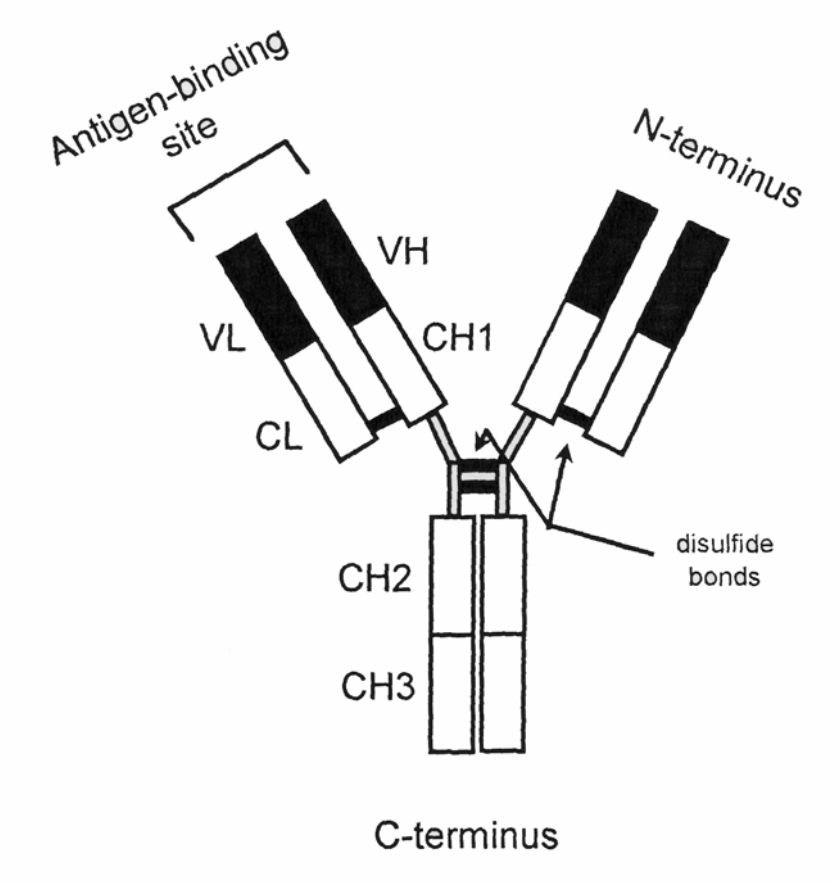
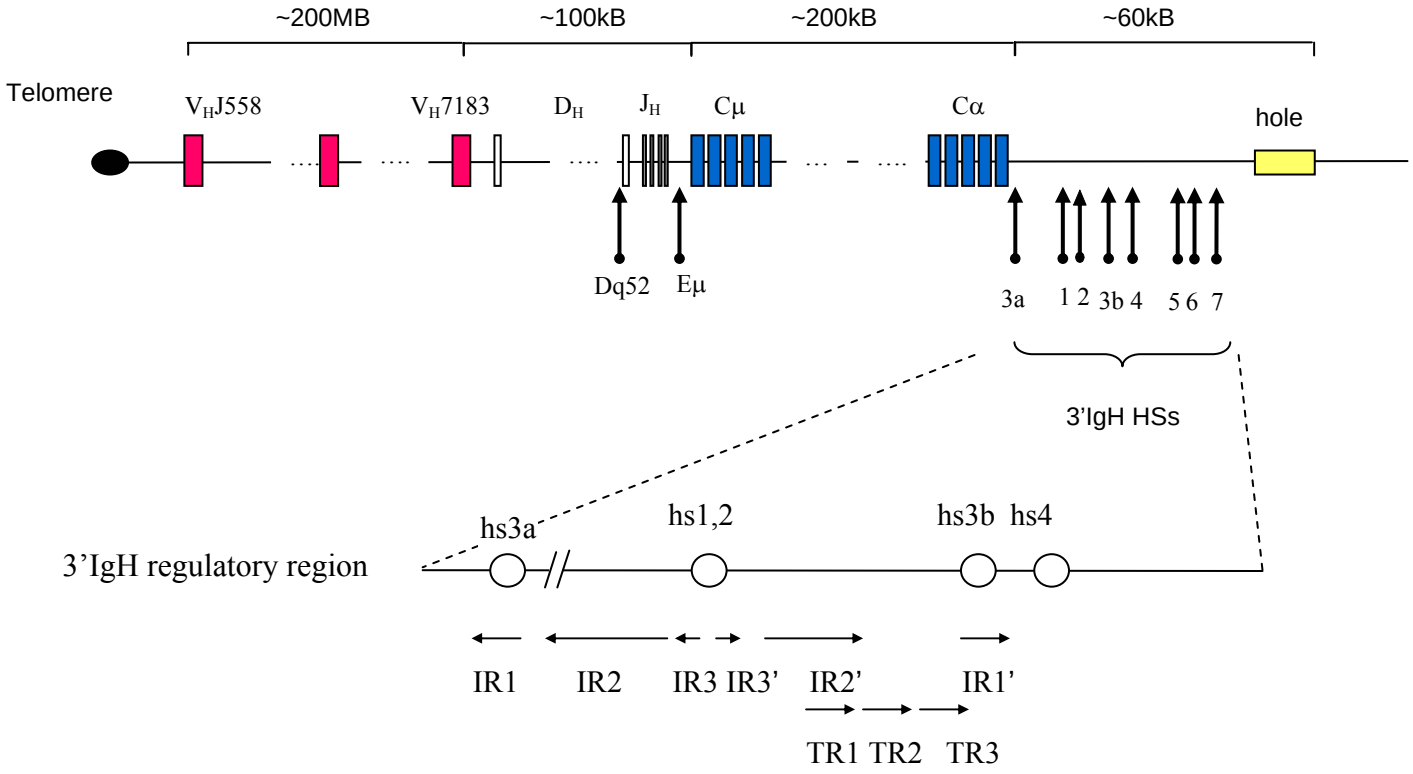


Figure 1-2 Organization of the IgH locus and the position of the hypersensitive sites. The IgH locus is organized into clusters of donor sequences: V_H clusters, D_H clusters, J_H clusters, and constant regions. The coding sequences are indicated by closed and open boxes. *hole* is the non-IgH gene next to the 3' end of the IgH locus. The span of each cluster is indicated by a bracket above the corresponding cluster. Vertical arrows point to hypersensitive sites. Horizontal arrows represent tandem repeats and inverted repeats. Open ovals represent enhancer elements. Tandem and inverted repeats are represented below the 3'IgH regulatory region (Chauveau and Cogne, 1996). TR: Tandem repeat; IR: Inverted repeat.

Figure 1-2 Organization of the IgH locus



Downstream of the J_H cluster and following a 7kb intron are 8 constant genes: Cμ, Cδ, Cγ3, Cγ1, Cγ2b, Cγ2a, Cε, and Cα, all separated by large introns. They are composed of three (α and δ) or four (μ, γ, and ε) exons, each encoding a functional unit of the extra-cellular domain of the Heavy chain, in addition to one or two membrane exons (Edelman et al, 1969). With the exception of Cδ, each constant region gene has its own so-called I exon promoter (IP). Transcripts initiated from these promoters are not translated into protein and therefore are termed "Sterile Transcripts". It is generally accepted, however, that sterile transcription initiated from these promoters is required for Class Switch Recombination (CSR) (Zhang et al., 1995). The next downstream gene is brain specific gene *hole*, which was mapped to 30kb downstream of hs4.

1.1.2 B cell development and the events that happen at the murine IgH locus

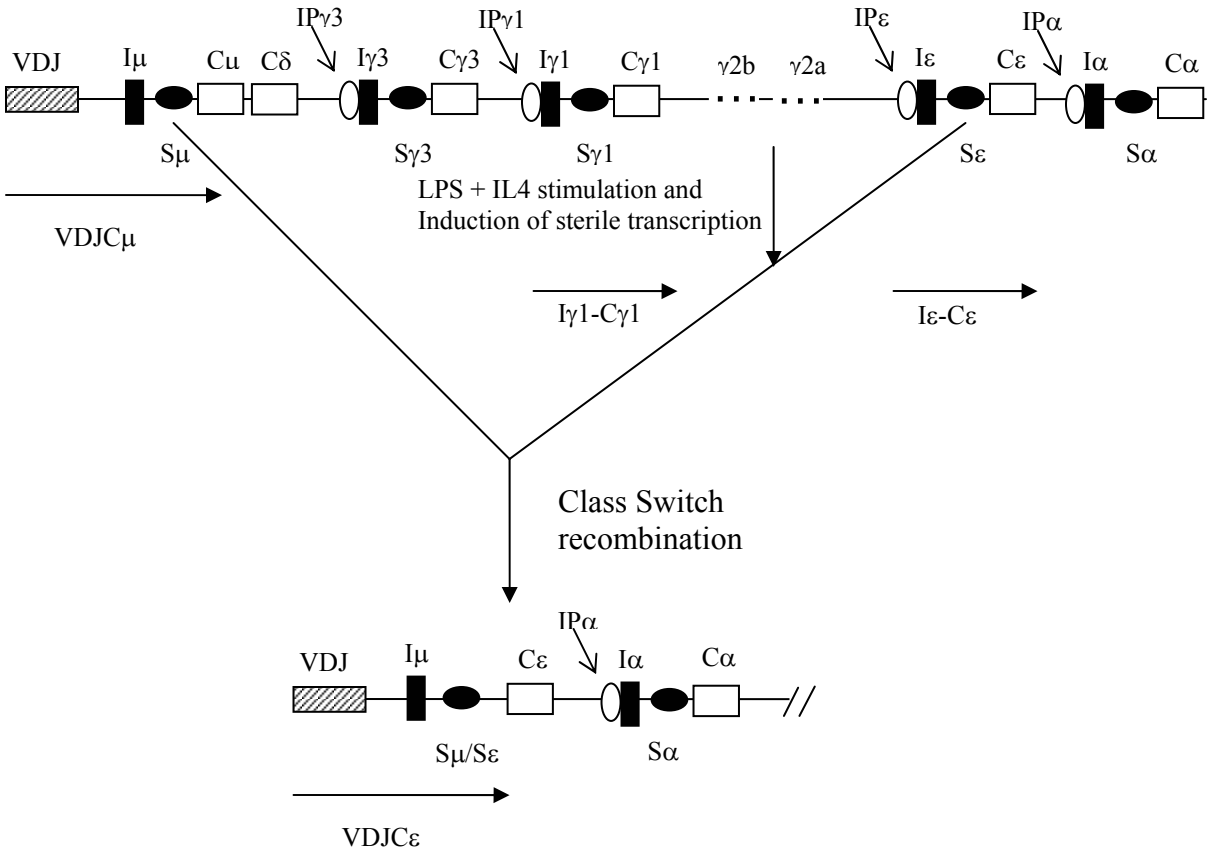
The antibody producing cells, B cells, start their journey from the bone marrow. Here, hematopoietic stem cells commit themselves to early pro-B cells, and develop all the way through immature B cells before they migrate to the periphery for further activation. During this process in the bone marrow, the heavy chain and light chain genes are assembled somatically through two rounds of DNA rearrangement. First, D to J joining occurs when early pro-B cells develop into late pro-B cells, followed by V to DJ joining to form a complete VDJ transcription unit, at which point the cells are designated large pre-B cells (Willerford et al., 1996). Heavy chain polypeptides expressed by the assembled VDJ gene upstream of Cμ are transported to the surface together with surrogate light chain consisting of two polypeptides: VpreB and λ5 to form the pre-B cell receptor. The pre-B cell receptor signals back to the pre-B cells to stop rearranging the other allele. Therefore, during B cell ontogeny, only one allele of the IgH locus has a

productive VDJ fragment assembled—a phenomenon called allelic exclusion (Reviewed in Mostoslavsky, et al., 2004). Next, the light chain gene is formed via V to J joining at the Ig light chain locus (either κ or λ locus), forming immature B cells. Then, the immature B cells are subject to scrutiny for having successfully rearranged the entire B cell receptor molecules before it enters the periphery, and become mature B cells on whose surface both IgM and IgD are displayed (Preud'homme et al., 2000). Why B cells should display both classes on its surface is presently unknown.

After successfully rearranging the variable regions of both heavy chain and light chain genes, mature B cells exit the bone marrow and, through the blood stream, migrate into secondary lymphoid tissues where a small subset of B cells are activated by stimulation by T cells in combination with antigens and cytokines. Depending on the kind of stimuli a B cell receives, it may undergo class switch recombination (CSR) to one of the heavy chain classes (Figure 1-3). In this process, the rearranged VDJ is juxtaposed to one of the six constant region genes downstream of Cδ, to generate antibodies of class IgG3, IgG2a, IgG2b, IgG1, IgE or IgA (Lorenz and Radbruch, 1996). During the activation process, mutation occurs with high frequency in the variable regions of both the heavy and light chain genes—a process called somatic hypermutation (SHM). The mutations in the variable region may result in null mutation, decrease in affinity, or increase in affinity to the antigen the B cell recognizes. Those B cells with high affinity to antigens are selected and clonally expanded to generate a cohort of B cells that recognize the same epitope of antigen. Since, overall, the affinity of the B cell pool to the invading antigen is increased, this process is also called affinity maturation from the perspective of serology (Reviewed in Neuberger et al., 2005).

Figure 1-3. Structure of constant region genes and mechanisms of class switch recombination. As representative, class switch recombination to IgE stimulated by LPS plus IL-4 and corresponding sterile transcription (represented by arrows) is shown. Hatched boxes represent rearranged VDJ, open boxes represent various C_H coding sequences, black boxes represent I exons, open ovals represent promoters of I exons (IP), black ovals represent various S_H regions. γ2b and γ2a constant region genes in the germline configuration of the IgH locus are omitted.

Figure 1-3 Structure of constant region genes and the mechanisms of Class Switch Recombination



1.2 General Mechanisms of Transcription Regulation

Three types of RNA polymerases-I, II, and III-are involved in transcription of eukaryotic genes. RNA polymerase I is responsible for the transcription of genes that encode ribosomal RNA; RNA polymerase II is responsible for the transcription of genes that encode mRNA; While RNA polymerase III transcribes genes that encode 5S RNA and tRNA. Since immunoglobulin genes are transcribed only as mRNA, we only introduce the mechanism of transcription by RNA polymerase II.

RNA polymerase II is a huge complex consisting of at least 12 subunits, one of which is a 220kD protein called the large subunit and which plays a central role in transcription. The C-terminal of the large subunit contains a repetitive seven amino acid sequence, subjected to phosphorylation at multiple sites. It is believed that such phosphorylation enables elongation through blockade of interactions between RNA polymerase II with pre-initiation complex (Peterson and Tjian, 1992).

1.2.1 Transcription initiation is the rate-limiting step of transcription.

Unlike prokaryotic RNA polymerase which binds to DNA directly as a template, eukaryotic RNA polymerases recognize only stable transcription complexes formed on the DNA template. Therefore, the rate-limiting step of eukaryotic gene transcription is the formation of a stable transcription complex with promoters (Hochheimer A and Tjian, 2003).

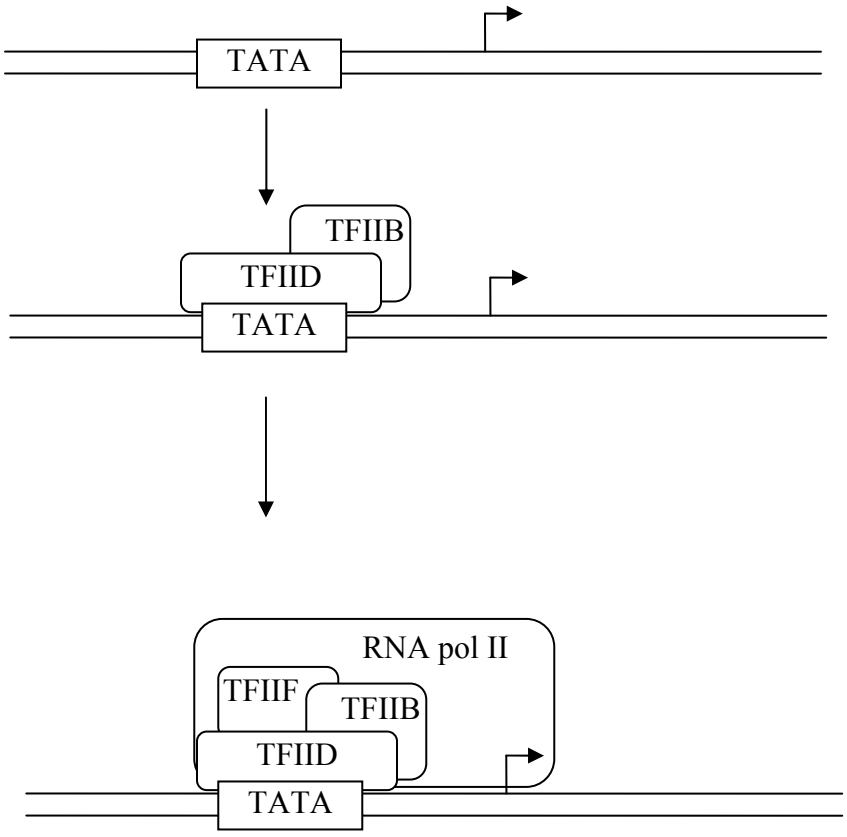
While the primary complex responsible for synthesizing mRNA from the DNA template is RNA polymerase II, the minimal complex for transcription initiation consists of TFIID, TFIIB, TFIIF and RNA polymerase II (Figure 1-4). RNA polymerase II is recruited to promoters by TFIIB and TFIIF. TFIID itself is a large protein complex

containing TBP (TATA box Binding Protein) and at least seven other proteins called TAFs (TBP Associated Factors) (Pugh and Tjian, 1990). TAFs make many kinds of contact with different transcription factors with different domain structures, such as SP1, NF1, and estrogen receptor. Of utmost interest is the finding that one of the TAFs, TAFII250, has histone acetyltransferase activity (Mizzen et al., 1996). As will be discussed later, histone acetylation is part of the mechanism to open chromatin conformation, which is now believed to be one of the earliest steps toward transcription activation. This finding suggests that open chromatin conformation might be an intrinsic function of transcription factors that are universally involved in transcription. TFIIB stabilizes the interaction between TFIID and DNA. The N-terminus of TFIIB is a zinc finger domain, oriented toward the transcription start site. It functions in recruiting both RNA polymerase II and TFIIF to the pre-initiation complex. The carboxyl terminus interacts with and stabilizes the complex of TBP and DNA (Ha et al., 1993). TFIIF is a two-subunit protein. It associates with RNA polymerase II even in the absence of DNA. It has functions in both transcription initiation and elongation (Price et al., 1989; Flores et al., 1992).

As discussed above, the assembly of a pre-initiation complex requires numerous protein-DNA, protein-protein interactions. This complexity provides many opportunities for modulation at multiple levels. Tissue specificity is achieved by combinatorial interactions between tissue specific and non-tissue specific transcription factors, although each interaction itself may not be sufficient to guarantee the specificity of the expression (Ernst and Smale, 1995).

Figure 1-4. Assembly of the pre-initiation complex at the typical promoter containing a TATA box transcribed by RNA polymerase II. Arrow represents the transcription start site.

Figure 1-4. Assembly of the transcription pre-initiation complex



These combinatorial interactions depend on the binding of tissue specific and non-specific transcription factors to cis-elements often found around DNaseI hypersensitive sites. As discussed in the next section, these cis-elements contribute to the activation of transcription from three aspects: opening chromatin, insulation from adjacent chromatin, and enhancement of transcription. These cis-elements, in addition to silencers, serve as hot spots for the binding of numerous nuclear proteins and often combine to form locus control regions (LCRs) which determine the correct tissue and stage specific expression of genes.

With regard to maintaining the activation state of a gene, tissue specific and developmentally regulated genes are also quite different from non-tissue specific genes. For example, transcription complexes that activate 5S RNA are stably bound to its promoter during transcription (Darby et al., 1998; Chipev and Wolffe 1992). However, the complexes are disrupted by progression of a replication fork (Wolffe 1993). By contrast, the passage of a replication fork did not have an effect on the binding of transcription factor complexes to the β-globin locus (Barton and Emerson, 1994). These results suggested that transcription factor complexes that activate non-tissue specific genes are disrupted and reassembled during each cell cycle, while those that activate tissue specific genes remain bound to promoters (and enhancers). The stable interaction of transcription factors with their binding sites may help to maintain chromatin configuration during cell proliferation and differentiation, which could serve as a molecular mechanism for the recent proposal that the IgH locus is activated in a stepwise manner (Chowdhury and Sen 2001).

Another mechanism by which maintaining chromatin configuration is carried out is the maintenance of the modification status of chromatin itself, which is called epigenetic imprinting (Reviewed by Delaval K and Feil R, 2004). During replication, nucleosomes on the old strand of DNA are displaced. Old histones, with preservation of their post-translational modification status, mix with newly synthesized histones and form hybrid nucleosomes. This may provide a means to replicate chromatin configuration from mother cells to daughter cells (Burhans et al., 1991). This idea is supported by the "histone code" theory which states that histone modification pattern can propagate along chromatin through a positive feedback mechanism (Jenuwein and Allis 2001).

1.2.2 Chromatin, DNaseI hypersensitive sites (HSs) and Locus Control Regions (LCR).

A simplified version of the current model for transcription activation is that the activation is accomplished in two steps: first, chromatin is shifted to an "open configuration"; Second, tissue- and stage- specific transcription factors are expressed and gain access to the chromatin and commence transcription. Therefore opening chromatin configuration is generally considered the prerequisite for transcription, although it is now believed that chromatin status is more complicated than the binary open-closed status (Gregory and Horz, 1998).

DNA in the nucleus is neither naked nor of linear superhelical shape. The double stranded DNA wraps around a protein complex consisting of core histones H2A, H2B, H3, and H4 to form a nucleosome. The nucleosome is the basic building unit in the formation of chromatin. Each nucleosome unit contains 146bp DNA. Nulceosomes are further packed by their association with the linker histone H1 to form 30nm fibers.

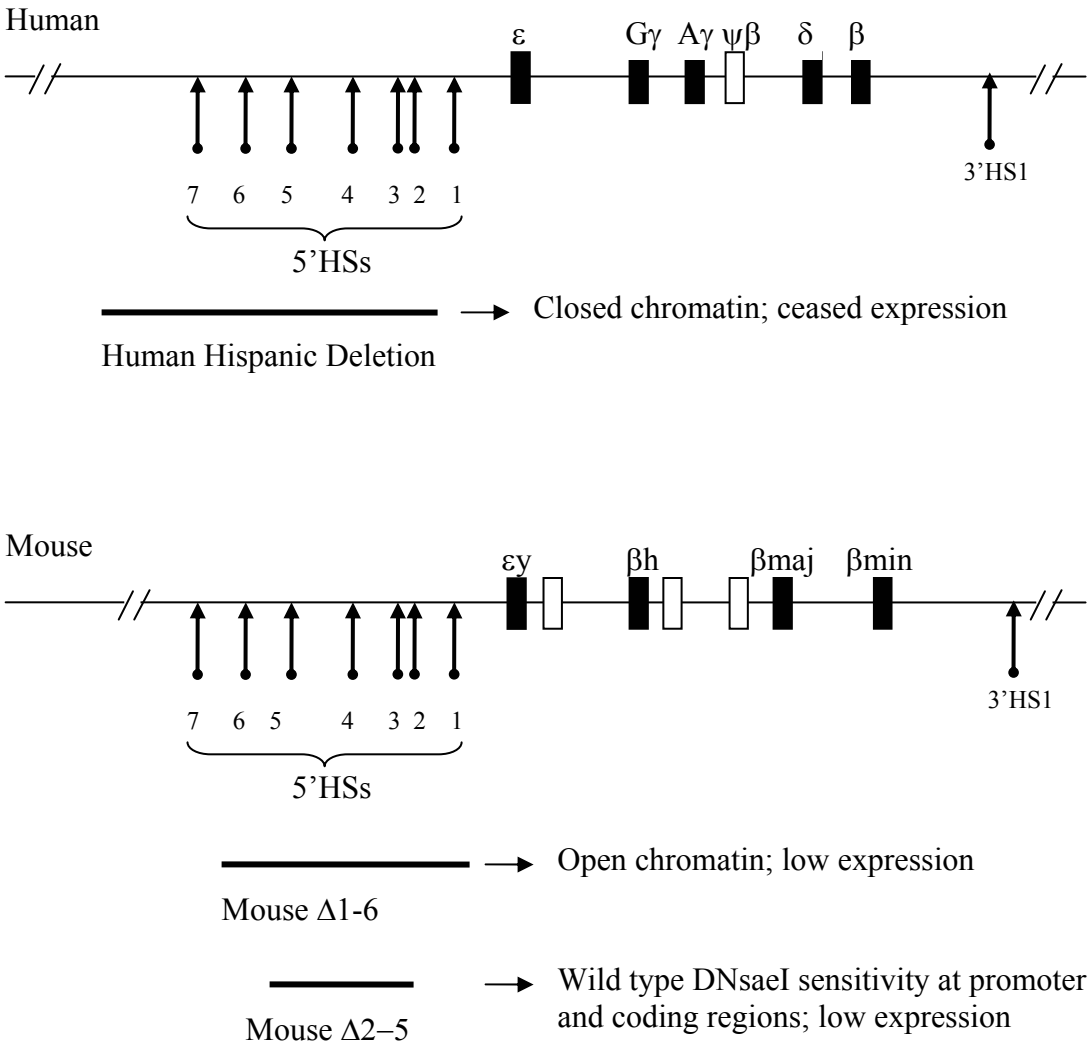
We currently know very little about chromatin structures higher than the 30nm fiber. In general, association of DNA with histones H1, H3, and H4 has a negative effect on transcription, presumably because histone H1 helps fold nucleosomes into the 30nm fiber, and the association of histones H3 and H4 with DNA is a prerequisite for histones H2A and H2B to bind DNA and form the nucleosome (Cremis and Yaniv, 1980).

As a measure of chromatin status, DNaseI has proven itself quite useful. A transcriptionally active gene usually is one order of magnitude more sensitive to DNaseI than the bulk of chromosomal DNA (Wood and Felsenfeld 1982). DNA sequences containing various kinds of cis-elements including: promoters, enhancers, insulators and the LCR in the narrow sense usually display sensitivity to DNaseI two or more orders of magnitude higher than the bulk of chromosome (Gross and Garrard, 1988).

The human β-globin locus has been regarded as a classic model system to study the interplay between chromatin and transcription (Figure 1-5). It is a multigene locus, which has five genes in the order from 5' to 3': 5'-ε-Gγ-Aγ-δ-β-3'; the genes are positioned in the same order as they are expressed during ontogeny. Seven DNaseI hypersensitive sites have been found to the 5' of the ε globin gene and one site has been found 3' to the β-globin gene. In the Hispanic Thalassemia, DNA sequences spanning HS2-HS5 and extending 25kb upstream of hs5 (including HS6 and HS7) are deleted, resulting in ceased β-globin expression. The deletion causes 5'HS1 to vanish and renders the whole locus resistant to DNaseI, resembling a "closed" chromatin configuration (Driscoll et al., 1989).

Figure 1-5 Human and mouse β-globin loci. The human and mouse β-globin loci each contain seven 5' DNaseI hypersensitive sites and one 3' DNaseI sites. Deletion of DNA sequences encompassing 5'HS2 through HS5 and extending 25kb upstream of HS5 (to include HS6 and HS7) in human causes Hispanic thalassemia and closed chromatin configuration in erythroid cells. Deletion of mouse HS1 through 6 or hs2-5 also caused a dramatic reduction in β-globin expression, but the chromatin structure was not greatly altered. Black boxes, coding sequences; open boxes, pseudogenes; arrows point to DNaseI hypersensitive sites.

Figure 1-5 Human and mouse β-globin loci



Transgenic mouse studies showed that a transgene linked to the mouse β-globin HS1~5 was expressed in a copy number dependent manner with the correct developmental pattern, suggesting that the transgene was expressed at the same level at any integration site (Grosveld et al.,1987). This result suggested that the β-globin HS1~5 contains all the elements required for opening chromatin, insulating the locus from the influence of neighboring chromatin, and for driving transcription in the right tissue at the right stage of development. Subsequently, HS1~5 was termed the β-globin locus control region (LCR) to indicate its omnipotent role in controlling transcription of the globin-like genes, although to date, two more DNaseI hypersensitive sites have been found 5' to HS1 through HS5. Later, copy-number dependent and position-independent expression of transgenes was used as a golden standard to define LCRs in other genes, such as the T cell receptor genes and Hoxb genes.

Although it is tempting to assume that the operationally defined LCRs contain all of the cis-elements that regulate both chromatin structure and expression of a tissue and stage-specific gene *in vivo*, further studies showed that this is not the case. 5'HS2~5 knock-out and 5' HS1~6 knock-out mice showed dramatically decreased expression of the globin genes, but the chromatin structure was not dramatically changed in terms of histone H3 and H4 acetylation pattern within the promoter region (5'HS2~5 knock-out) nor in terms of general DNaseI sensitivity of the whole locus (5'HS1~6 knock-out) (Schubeler et al., 2001; Epner et al, 1998; Bender et al., 2000). These results, combined with the observed phenotype of the Hispanic Thalassemia, suggest that there exist control

elements 5' of hs6 (in the murine locus) that serve a redundant role with the LCR (hs1-5) with respect to opening chromatin structure (Figure 1-5).

This introduces the possibility that some unknown cis-element 5' to HS6 could exert the same function as some elements inside the 5'HS1-6 region and that the combination of these unknown cis-elements with part of the DNA contained within HS1-6 could constitute another LCR.

1.3 Mechanism of transcription activation in the IgH locus

1.3.1 Transcription factors involved in transcription within the IgH locus

Cis-elements in the IgH locus include various kinds of promoters, five known enhancers, matrix attachment regions and at least eight DNaseI hypersensitive sites (HSs). Six of these HSs had been discovered when this study began. They are, in the 5' to 3' order, Eμ, hs3a, hs1,2, hs3b and hs4. All of them have transcription enhancement activity. Recently, three more HSs were found downstream of hs4, namely, hs5, hs6, and hs7, which do not have enhancer activity but may function as an insulator. (Garrett et al., 2005)

Eμ and the 3' IgH enhancer region (hs3a, hs1,2, hs3b, hs4) contain a series of transcription factor binding sites. These include the octamer binding site, a series of E box elements μE1-5, NF-κB binding sites and BSAP binding sites, as summarized in Table I. Study of the effects of transcription factors which bind to the cis-elements in the IgH locus, especially those shown to be involved in VDJ recombination, Class Switch Recombination as well as high level expression, has and will continue to contribute to explaining the function of enhancers in these processes.

Table I Protein binding motifs in the enhancers in the IgH locus.

Enhancers	Size (Restriction fragment)	Position. Kb downstream of $C\mu_m$	Protein binding motifs
Eμ	1kb	Between J_H and Cμ	μE1~5[*], Oct[*], μA[*], μB[f,*]
HS3a	447bp (StyI—AccI)	4.5 kb downstream of $C\alpha_m$	Oct[*], μE5, AP-1, AP-2, AP-4, kE2[a], MARE*
HS12	596bp (StuI-EcoRV)	12.5 kb downstream of $C\alpha_m$	Oct[*], μE1, μE5, AP-1, AP-2, AP-4, BSAP [b,d,*], NF-κB[*]
HS3b	1182bp (XbaI--Sau3A)	26 kb downstream of $C\alpha_m$	Oct*, μE3, μE5, AP-1, AP-2, AP-4[c], MARE[e,*], YY1[g]
HS4	1381bp (PstI—HindIII)	30 kb downstream of $C\alpha_m$	Oct*, μE3, μE5, AP-1, AP-2, AP-4, BSAP[*,c,d], NF-κB*

Summarized from a. (Matthias and Baltimore, 1993) b. (Lieberson et al., 1991). c (Madisen and Groudine, 1994) d. (Michaelson et al., 1996). e (Muto et al., 1998). f. (Ernst and Smale.,1995); g. (Gordon et al., 2003.)

* Confirmed *in vitro* by EMSA assays.

None of these transcription factors is strictly B cell specific. Rather, some like NF-κB and Oct1, are quite widely expressed. The tissue and stage specificity of the element is probably achieved by the combination/cooperative binding of these transcription factors (Ernst and Smale 1995).

E-Box binding proteins: μE1-5 are a series of elements recognized by a group of helix-loop-helix transcription factors (bHLH). These bHLH proteins bind to these E-Box elements as homodimers or heterodimers in a B cell specific manner, although they are ubiquitously expressed (Eckhardt, 1992). These proteins can be divided into two classes. The first includes protein products of E2A, E2-2, and HEB genes, which bind μE2, μE4, and μE5 sites (Hu et al., 1992). The second includes protein products of TFE3, TFEB, and USF genes, which bind μE3 motif in Eμ and some IgH promoters (Artandi et al., 1994). This class of E-box proteins are characterized by the presence of a second protein-protein interaction motif—a leucine zipper (Kerppola and Curran, 1991). The zipper domain seems to determine the specificity of dimerization. The presence of two protein-protein interaction domains might enable these classes of bHLH proteins to form higher order multimers that would mediate interactions of promoters and distant enhancers (Artandi et al., 1994).

NF-κB: NF-κB is involved in regulating the expression of numerous genes and many cellular processes. Its importance in the immune response is exemplified by its involvement in the developmental processes, cellular growth, activation and apoptosis of lymphocytes (Li, et al., 2002a). It has five family members: Rel (c-rel), RelA (p65), RelB, NF-κB1 (p105/50), and NF-κB2 (p100/52) (Reviewed in Nolan and Baltimore, 1992). Its activation is mediated by the phosphorylation of its inhibitor IκB, resulting in its

translocation into the nucleus. Study with $p50^{-/-}$ mice showed that NF-κB is critical in the activation of sterile transcripts initiated upstream of C_H genes and for class switching. It has binding sites in both hs4 and hs1,2. It has been shown, by transient transfection assays, that NF-κB is involved in stage-specific regulation of hs1,2 and hs4. (Discussed in more detail below) (Michaelson et al., 1996; Snapper et al., 1996)

BSAP: As the master regulator of B cell development, BSAP regulates numerous genes involved in hematopoietic cell development. One of the most important BSAP functions is to suppress transcription factors that would promote the development of cell lineages in blood other than the B lineage. The other important function is to maintain B cell identity, including suppressing the expression of the numerous proteins that are highly expressed in Ig secreting cells, until B cells develop into plasma cells, at which point BSAP expression ceases (Schebesta et al., 2002).

Oct1/Oct2/OcaB: The octamer binding site, 5' ATGCAAAT-3', is present in both promoters and enhancers in the IgH locus. It is recognized by both the ubiquitous protein Oct1 and the B cell specific factor Oct2. OCA-B is a transcription co-factor that interacts directly with Oct1 and Oct2 (Luo et al., 1992; Luo and Roeder, 1995). It w as shown in transient transfection assay and transgenic mouse studies that the Oct-2/OCAB complex is critical in the activation of mature B cell and for class switching, but it has little effect on pre-B cells and plasmacytoma (Kim et al., 1996; Nielsen et al., 1996; Qin et al., 1998; Schubart et al., 1996; Tang and Sharp, 1999) Oct1 and Oct2 are POU domain proteins. The POU domain was first found for its ability to direct correct body sections during development (Herr et al, 1988). Octamer binding sites are found in various places in the mouse IgH locus, including V_H promoter, hs1,2, and hs4. Although the existence of an

octamer binding site in every V_H promoter is critical for these promoters' B lineage restricted acitivity, promoters of non-B cell specific genes also possess an octamer binding site. It was thought that this paradox might be explained if the IgH associated-octamer was only able to bind Oct2 and/or Oct2/OcaB heterodimer found only in B-lineage cells. However, B lineage cells in Oct2 and OcaB conditionally double knockout mice were able to develop all the way through the surface Ig positive stage. Since these mice lack germinal centers, it is not known, given the right environment, whether these B lineage cells could potentially develop into plasma cells and express the secreted form of heavy chain (Schubart et al., 2001). One of the many pieces of evidence suggesting specific roles for Oct2 in IgH expression is the fusion study done in our lab, in which only Oct2 could rescue the B cell characteristics, including expression of heavy chain in B x T fusions (Radomska et al., 1994). This ability of Oct2 to rescue B cell gene expression has been attributed to the C-terminus of Oct2 (Sharif et al., 2000).

PU.1. PU.1 belongs to the ets family, which is characterized by a conserved 84 amino-acid "ets" DNA binding domain, rich in basic and leucine residues (Nye et al., 1992). It is involved in the development of multiple hematopoietic lineages (Scott et al., 1994). Its involvement in early B cell development is manifested by its control over IL-7 receptor α expression and thus IL-7 dependent proliferation and differentiation (Dekoter, et al., 2002). PU.1 has been shown to bind μA and μB motifs in Eμ and 3'IgH enhancers *in vitro*. Ectopic expression of PU.1, and another ets family protein, ets-1, was sufficient to activate Eμ core in non-lymphoid cells (Nelsen et al., 1993). PU.1 increased chromatin accessibility to Eμ core enhancer in pro-T cells, suggesting that PU.1 has the ability to influence chromatin structure (Nikolajczyk et al., 1999). Further analysis showed that

PU.1 is able to increase accessibility to sites proximal to its binding site but not widespread chromatin accessibility at the IgH heavy chain locus (Marecki et al., 2004).

BSAP, NF-κB and octamer binding sites were shown to have positive and/or repressive effects on hs1,2 and hs4 function in transient transfection assays (Michaelson et al., 1996). EMSA, combined with mutation analysis on these binding motifs in transient tranfections, showed that different enhancers in the same stage of B cell development could be differentially regulated by the same (set) of transcription factors. In mouse, at the surface Ig positive B cell stage, Oct1, BSAP and an NF-κB-like complex contribute to the enhancer activity of hs4 but repress hs1,2 activity. At the plasma cell stage, Oct1 binding does not affect the activity of hs4, but enhances hs1,2 activity. This difference could be due to the fact that the transcription factor binding sites in hs1,2 and hs4 are placed in different order and at different distances. Another possibility is that there are differences in the interactions of other transcription factors with these two enhancers that would affect the outcome of binding of this common set of transcription factors (Michaelson et al., 1996).

It was also shown that the same (set of) transcription factors could have different effects on the same enhancer element at different stages of B cell development: BSAP suppresses hs4 activity at the pre-B cell stage, but has no effect or even enhances hs4 activity in B cells (Michaelson et al., 1996). This may result from the temporal change in BSAP expression level in these stages. There are two BSAP binding sites in hs4, one with high affinity and the other with low affinity. At the pre-B cell stage, BSAP is expressed at high levels and both sites are occupied. At the B cell stage, BSAP is less abundant, and therefore, only the site with high affinity to BSAP is occupied. In the case

of NF-κB, the different effect at different developmental stages could result from the complexity of the NF-κB complex itself. Four of the five NF-κB family members have been shown to participate in the regulation of the 3' IgH enhancers. The complex that binds to hs1,2 at the surface Ig positive B cell stage involves p50 and C-Rel, while the complex binding hs1,2 at the plasma cell stage involves p50, p52, and RelB (Michaelson 1996). In summary, there are complex interactions taking place at the individual enhancers, and, in the actual IgH locus, all of the enhancers are together in the chromosome. This calls for studying enhancers not in isolation but in their natural positions and orientations, taking into account chromatin structure effects.

1.3.2 Functional Studies of individual IgH cis-elements.

All known Cis-elements in the IgH locus fall into three categories: promoters, enhancers and insulators. The promoters can be further divided into two groups: V_H promoters which give rise to transcripts that are translated into peptides that constitute the heavy chain of B cell receptors or antibodies; I exon promoters (IP) that give rise to sterile transcripts that are not translated into peptides. The exact molecular mechanism of how these elements are differentially regulated is not clear. It can be speculated that the availability of transcription factors and the configuration of chromatin are important.

Eμ, hs3a, hs1,2, hs3b, and hs4 have been shown to have transcriptional enhancement activity by themselves in certain stages of B cell development. Eμ possesses chromatin opening ability (Jenuwein et al., 1997). Expression of hs1,2-linked genes in transgenic mice suggests that hs1,2 is able to open chromatin by itself (Arulampalam et al., 1994a; Arulampalam et al., 1994b; Andersson et al., 1999). However, none of the other three 3'IgH enhancers has been tested alone for their ability

to influence chromatin structure in transgenic mice. It is possible that the synergies and redundancies, which can only be tested by experiments involving different combinations of these cis- regulatory elements, play more important roles in the regulation of chromatin structure and transcription.

However, in the β-globin locus, specific functions have been assigned to individual elements, or even subdomains of individual elements (Recillars-Targa et al., 2002; reviewed in Li et al., 2002b). This dissection of the β-globin LCR seems to argue that there is a clear division of labor among these cis-regulatory elements. Therefore, functions of individual elements are first discussed, while the synergies and redundancies among them are discussed under "Enhancers in Combination".

V_H Promoters

Promoters in the IgH locus include V_H promoters and promoters for sterile transcription upstream of the coding sequences for each constant region. The V_H promoters, pV_H, all contain an octamer motif, TATA box, and CAATT box. In the case of the promoter of V_HB1-8—the V_H gene used in constructing the various BACs in this study, there are two transcription start sites (Neuberger, 1983). V_H promoters are mostly lymphoid specific, which contributes to the tissue specificity of transcription in the IgH locus.

DQ52 promoter

This is the most 5' HS found in the IgH locus. Sterile transcript μ^o is initiated here and is thought to regulate D_H J_H joining. In one Eμ knockout study, the lack of μ^o germ line transcript was correlated to lack of V-DJ joining on the cis-chromosome while DJ

joining occurred normally (Chen et al., 1993). It has also been shown recently that although deletion of the DQ52 promoter did not abolish D-J joining, it changed the J_H usage in D-J joining (Nitschke et al., 2001). If we assume that the 3'IgH regulatory region does not become active until the surface Ig positive B cell stage (discussed below), these results suggested that DQ52 participates in D_H to J_H joining.

Eμ

Eμ lies within a 1kb XbaI fragment located between the last J segments and the first constant segment Cμ (Figure 1-2). It is the first enhancer found in this locus for its enhancement effect on reporter genes (Banerji et al, 1983; Gillies et al, 1983; Neuberger, 1983). It contains a 350bps core (denoted cEμ hereafter) with a constellation of transcription factor binding sites (reviewed in Eckhardt, 1992), flanked by two matrix attachment regions (MARs) that bind to several matrix-derived proteins (Oancea et al., 1997). It was shown by knockout mouse studies that Eμ is required for efficient V-DJ assembly (Chen et al., 1993; Serwe and Sablitzky, 1993). Eμ is sufficient to drive transgene expression in transgenic mice in most integration sites. MARs that flank cEμ are required for this LCR like activity (Forrester et al., 1994). It was also shown that the chromatin opening activity of Eμ does not depend on transcription from the core (Jenuwein et al., 1997). Although transcription activation has always been correlated with chromatin opening, it has not been proven absolutely that chromatin opening is a prerequisite for activation of transcription.

hs3a and hs3b

hs3a and hs3b were discovered by mapping enhancer activity in the 3' end of the IgH locus and by DNaseI hypersensitive assays combined with transient transfection assays, respectively (Matthias and Baltimore 1993; Madisen and Groudine 1994). They are at the two ends of a palindrome sequence and are almost identical (95% identity) in sequence, but with opposite orientation (Saleque et al., 1997; Chauveau et al., 1996). Deletion of hs3a did not have any marked phenotype in mice (Manis et al., 1998). EMSA assays showed that hs3b binds transcription factors Oct1, Oct2, Pax5, and YY1 in a B cell line following LPS stimulation. It also binds NF-κB in a plasmacytoma cell line (Gordon et al., 2003). Transient transfections showed that they are not active when assayed alone until the B cells are activated, and even then they are weak enhancers. However, there is synergistic activity between these two weak enhancers in plasmacytoma cell lines (Saleque et al., 1997; Ong et al., 1998). It was shown in transient transfection assays that hs3b synergizes with hs4 at the surface Ig positive B cell and plasma cell stages (Ong et al., 1998)

hs1,2

Mouse hs1,2 was discovered by screening a phage DNA library for enhancer activity, and cloning the mouse counterpart for Rat hs1,2 (Dariavach et al., 1991; Lieberson et al., 1991). Although single deletion of hs1,2 has not yielded any discernable defect in mouse (Manis et al., 1998), since it was the first 3'IgH enhancer to be discovered, hs1,2 is the most extensively studied of these enhancers so far. Transient transfection assays showed that hs1,2 alone is not active until B cell develop to the plasma cell stage. However, it might cooperate with other 3'IgH enhancers (Ong 1998). It

was also shown that the activity of hs1,2 is elevated upon B cell stimulation by cytokines and that hs1,2 can synergize with the sterile promoters of Cε and Cγ2b, suggesting that hs1,2 might contribute to sterile transcription and CSR (Laurencikiene et al, 2001). An enhancer knockout mouse study showed that hs1,2 is not required for class switch recombination when other parts of the IgH locus remain intact (Manis, et al. 1998). hs1,2 was implicated in the high level expression of IgH genes in plasma cells in an enhancer knockout study in an Ig-secreting cell line (Lieberson et al., 1995), but as suggested in that study, replacement with a neo^r gene may have diverted the function of other enhancers upstream and downstream of hs1,2.

hs1,2 contains at least six protein binding sites. They are the octamer, an NF-κB site , two BSAP sites, one G-rich region, and PU.1 binding sites. In surface Ig positive B cells, the octamer and G-rich region have minimal activation activity on their own, but they synergize when both are present. In plasma cells, the effects of these two sites are additive by mutation analyses (Singh and Birshtein, 1996). The inactivity of hs1,2 by itself prior to the plasma cell stage can largely be attributed to repression by the two BSAP binding sites. The 5' BSAP binding site—the "a" site, is a strong binding site. The 3'BSAP binding site—the "b" site, is a weaker binding site. Both BSAP sites must be present for full repression. Interestingly, single mutations of the octamer, the NF-κB binding site, and the G-rich protein-binding site all led to relief of repression of hs1,2 by BSAP (Michaelson et al., 1996; Singh and Birhstein 1996). It was also shown that PU.1 cooperates with BSAP to repress hs1,2 by recruiting Grg4 (Linderson et al., 2004). These results suggest that presence of all of these binding sites is required for BSAP to exert its repressive activity. So far, hs1,2 is the only regulatory element found to be suppressed by

BSAP. Apparently, the repression of hs1,2 by BSAP is achieved only by concerted action of these transcription factors.

hs4

hs4 was discovered by mapping DNaseI hypersensitive sites downstream of Cα and is the strongest transcriptional enhancer found in the 3' IgH regulatory region (Madisen and Groudine 1994; Michaelson et al., 1995). In addition, hs4 distinguishes itself from the other 3'enhancers by displaying DNaseI hypersensitivity in pro-B and pre-B cells, suggesting a possible role in VDJ recombination. Four putative protein binding sites have been shown to bind the corresponding transcription factors by EMSA (Michaelson et al, 1996). They are, from 5' to 3', octamer, NF-κB, BSAP, and another octamer binding site. The octamer site binds Oct1 instead of Oct2 in surface Ig positive B cells, which might offer an explanation for the phenotype of $Oct2^{-/-}/OcaB^{-/-}$ mice (Michaelson et al., 1996; Schubart et al., 2001). In contrast to mouse hs4, human hs4 has only two of the four protein binding sites mentioned above: octamer and NF-κB. Interestingly, unlike the octamer site in mouse hs4 which binds only Oct1 by EMSA, human hs4 binds to both Oct1 and Oct2 in plasmacytomas, suggesting possible differences in roles played by the two Oct proteins (Sepulveda et al., 2004).

Enhancers in combination

Assaying activities of enhancers alone might not provide an accurate account of the function of the enhancers. For example, hs1,2 did not display much enhancer activity in transient transfection assays, but it might cooperate with other enhancers to have an additive or synergistic effect with other 3'IgH enhancers (Ong et al., 1998). Because of

the existence of synergies and redundancy within the 3' IgH enhancers, it might be more informative to examine them in pairs. hs3a & hs1,2 pair deletion from a mini-cassette did not have any effect on the already established expression in B or plasmacytoma cell lines, while hs3b & hs4 pair deletion had dramatically decreased transcription from the same construct in cell lines representing the plasma cell stage but not in cell lines representing the surface Ig positive stage (Shi and Eckhardt, 2001). From the perspective of an established environment, this result suggests that the hs3a & hs1,2 and hs3b & hs4 pairs, respectively, are able to compensate for the loss of one another in surface Ig positive cells, but the function of hs3b and hs4 is unique in plasma cells and cannot be replaced. Deletion of the hs3b & hs4 pair in mice led to a defect in CSR to every heavy chain class except IgG1 (Pinaud et al., 2001). In contrast to the study in cell lines mentioned above in which the function of hs3b & hs4 could be compensated by hs1,2 & hs3a at the surface Ig positive stage, this result points out the unique attributes of the hs3b and hs4 pair in directing CSR even in the presence of Eμ, once cells are stimulated to undergo CSR. The contrast could reflect the difference in the mechanism of transcription from a V_H promoter and sterile transcription from the intronic promoters (Ip) upstream of each heavy chain constant region (C_H). Another possibility is that the activation of CSR is a transient process which cannot be tested in cell lines.

In summary, it has been established that Eμ plays a pivotal role in early B cell development. hs4 starts to show enhancer activity and DNaseI hypersensitivity at the pre-B cell stage, while hs3a, hs1,2, and hs3b do not show DNaseI or enhancer activity until the surface Ig positive stage of B cell development (Lieberson et al., 1991; Gianini et al., 1993., Michaelson et al., 1995; Ong et al., 1998) . It can be postulated from the transient

transfection assays that, during the process of B cell development, 3'IgH enhancers play a more and more important role in transcription and recombination events. It has firmly been established that Eμ is required for V_H to D_HJ_H recombination, while the pair hs3b and hs4 is required for CSR in the presence of Eμ.

A most notable feature of the 3' enhancer region is its palindromic structure. As shown in Figure 1-2, there are at least three large inverted repeats, named IR1/IR1', IR2/IR2' and IR3/IR3'. hs3a and hs3b are symmetrically distributed into IR1 and IR1'. Consequently, many transcription factor binding motifs are symmetric, a feature shared by IR2/IR2' repeats. Part of IR2' is also a tandem repeat TR1, with TR2 and TR3 lying successively downstream. The symmetric axis of IR3/IR3' is hs1,2, with a 0.6kb purine rich region lying downstream. There are also three Z-DNA tracts which are regions of uninterrupted alternating purine and pyrimidines (Chauveau et al., 1996).

Others have suggested that a palindromic structure is an indication of a locus control region (LCR) (Chauveau and Cogne., 1996). There are two 0.6kb tandem repeats flanking the hs3 element of the murine β-globin LCR and a combination of inverted and tandem Alu repeats scattered in the human adenosine deaminase gene LCR, implicating the transposon origin of 3' enhancers. (Chaveau and Cogne, 1996; Mills et al., 1997). It has been proposed that these palindromic structures alter chromatin structure, thereby conferring its effect on transcription. Based on the postulated higher-order DNA structure, it was proposed that the 3'IgH regulatory region can be divided into two functional units. Unit I is comprised of DNA sequences spanning hs3a through hs3b. Unit II is comprised of hs4 (Saleque et al., 1997). Taking into consideration the recently discovered hs5 through hs7, Unit II might be comprised of hs4 through hs7 (Garrett et al, 2005). In

support of this hypothesis, it has recently been shown by sequence analysis that the higher-order DNA structure, or "switch-like" structure, is well-conserved from rodent to human, rather than the primary DNA sequence of the 3'IgH regulatory region (Sepulveda et al., 2005). The physiological relevance of this division, and the division of labor among these postulated functional units need to be evaluated by further experimentation.

1.3.3 The locus control activity of the 3'IgH regulatory region

The physiological relevance of the regulatory activity of the 3'IgH region can be exemplified by the existence of several kinds of lymphomas that have arisen from the translocation of genes involved in cell cycle and differentiation to the IgH locus. These genes are: Myc, Bcl2, and Pax 5 (Madisen and Groudine, 1994; Cazzaniga et al., 2001; Rack et al., 1998). All of these translocations result in juxtaposition of the proto-oncogenes with the 3'IgH regulatory region. These results not only demonstrated that Eμ is dispensable for high level gene expression but also showed that the 3'IgH regulatory region is able to maintain an open chromatin configuration when placed in close proximity with various types of chromatin.

In addition, the tissue specific and cell-stage specific DNaseI hypersensitivity of the 3' IgH enhancer region promoted scientists to address the question whether the 3' enhancer region acted as a locus control region (LCR). An LCR has been described as a series of cis-elements that separate the gene under its control from the adjacent chromosome environment and regulate the gene temporally and spatially. LCRs could fall into two categories based on the variability of transgene expression. The first LCR discovered is the β-globin LCR (Grosveld et al., 1997). When transgenes under the control of the LCR were integrated into random sites of the genome, the transcription

level was dependent on copy number. Later LCRs of other loci were discovered, but most of them did not confer transgene expression in a strictly copy number-dependent manner, although they could ensure expression in most integration sites (summarized in Li et al., 2002b). To distinguish the two standards and for the convenience of discussion, the β-globin like LCR has been called an **LCR**, but the less strict LCRs **partial LCRs** (Ellis et al., 1996).

Previous studies are in agreement with the notion that the IgH3' regulatory region is a partial LCR, that is, stably transfected genes and transgenes in mice under the control of the full complement of 3' enhancers are exclusively expressed in B cells (Arulampalam et al., 1994b; Chauveau et al., 1999). However, whether the 3' enhancer region has full or "complete" LCR activity is under debate. A stably transfected *c-myc* gene under the control of all the IgH 3'enhancers except hs3a was shown to be expressed in a position-independent, copy number-dependent manner in 1165, a plasmacytoma cell line (Madisen and Groudine, 1994). In contrast, a human β-globin gene carrying all four 3'IgH enhancers in transgenic mice did not display copy number-dependent expression (Chauveau et al., 1999). This discrepancy may result from the incomplete enhancer region in the constructs in both studies. The former lacked the hs3a region and neither had a complete inverted repeat flanking hs1,2, which was shown to have a significant effect on the function of hs1,2 (Chauveau et al., 1999). In addition, the promoters in these two studies were quite different, the former being both V_λ and *c-myc* promoters while the latter study used a V_H promoter.

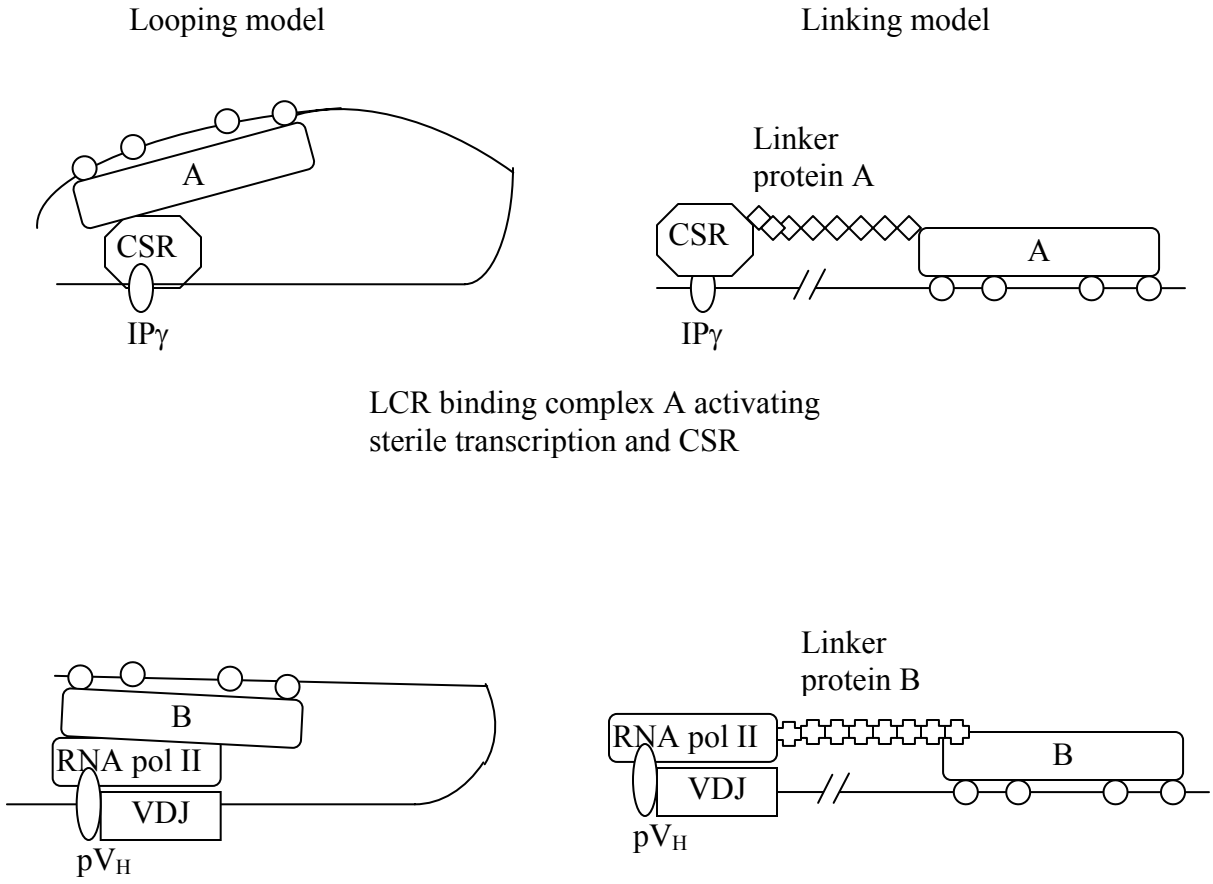
According to the functions originally attributed to an LCR, it should at least contain the following three elements: (1) Chromatin conformation modulators that open

the chromatin and keep this open configuration in the right tissue and at the right stage of development and that close the chromatin and keep the transcription factors or recombination machinery inaccessible to the locus in other inappropriate tissues; (2) Expression-level regulators including tissue- and stage- specific enhancers and silencers. (3) Insulators that keep the locus free from the influence of adjacent chromatin. We can postulate that a partial LCR has the first two kinds of elements so that the transgene under its control can be expressed in the right cell at the right time. However, because it lacks the insulator, the expression level is under the influence of elements in the adjacent chromatin. There are two possibilities considering the IgH LCR. First, there is an LCR in the IgH locus but all the elements required for its function have not yet been identified or not constructed. Second, the IgH locus is regulated by only a partial LCR and the reason for the locus to function normally without an insulator is that, by chance, there are no regulatory elements (especially enhancers and silencers) surrounding the locus that are able to influence IgH expression and/or recombination. The notion that an enhancer or silencer can function over a long distance, in the case of the IgH locus, more than 1Mb, suggests that it is more likely that the IgH locus has an LCR but that the entire LCR has not yet been defined.

In the IgH locus, as well as in the β-globin locus, the function of the LCR is more complicated than that of other LCRs such as the LCR of the human CD2 gene. There are a series of gene segments in the locus and each has its own promoter. It has been widely accepted that expression of one gene in the locus represses that of the other (Giglioni et al., 1984).

Figure 1-6 Looping and linking models in the IgH locus. Two events on the IgH locus are illustrated. Class switch recombination, and transcription from the V_H promoter. Both require different protein complexes to bind the LCR. The linking model requires linker proteins consisting of repeats of units of protein complexes.

Figure 1-6 Looping and linking models in the IgH locus



The LCR, in cooperation with transcription factors, should determine which promoter initiates transcription at a certain stage of development working over a long distance. Two general models have been proposed to explain how the cells manage to have the right promoter interact with the LCR (Figure 1-6). One of them is the looping model (Dunaway and Droge, 1989; Mueller-Storm et al., 1989). Promoters with different transcription factor binding sites in different positions in the same locus as the LCR compete to interact with the activated LCR. Only the promoter interacting with the activated LCR elements can serve as a site for initiation of transcription. The distance between a certain promoter and the enhancer functioning at the time is a critical parameter in determining the transcription efficiency from this promoter. A proximal promoter has privilege over a distant promoter and is more efficiently transcribed. (Reviewed by Dillon et al., 1997). Meanwhile, the interactions between transcription factors and promoters as well as other elements in the LCR alter the priority of access of promoter to LCR, as illustrated by altered gene expression profiles of β-globin like genes in yolk-sac stage but not in the adult stage (Tanimoto et al, 1999). This interaction is determined by both the abundance of the transcription factors and the DNA modification state. For instance, methylation in an insulator results in expression of a distant gene and silencing of a proximal gene (Hark et al., 2000). An alternative model is the linking model, which states that transcription factors bind at both the LCR and promoters, forming several docking sites. Non-DNA-binding facilitator proteins form a continuous protein chain bridging the docking site on the LCR and one of the docking sites on promoters (Morcillo et al., 1997; Engel and Tanimoto, 2000; Milot et al., 1996).

The distinctions between these two models are: 1. Topology: the looping model requires that DNA or chromatin loop back to cover the distance between the LCR and promoters while in the linking model, the bridge formed by facilitator protein chains cover the distance between LCR and promoters. 2. The composition of protein complexes that connect the LCR and promoters: in the looping model, the protein complexes that connect the LCR to one of the promoters can be quite different from those that connect the same LCR to another promoter. In the linking model, it is very unlikely that different promoters use different facilitator proteins. Otherwise, the IgH locus would require hundreds of kinds of facilitator proteins to facilitate the hundreds of different promoters. In addition, the composition of each protein chain would probably consist of uniform repeats of a protein complex, analogous to the composition of actin.

The complexity of the events that take place in the IgH locus requires variations of the molecular functions exerted by the protein complexes that bind LCR elements, which is much more easily achieved by the looping model than by the linking model (Figure 1-6). i.e. Class switch recombination and transcription from the V_H promoter would require different molecular machineries assembled at both LCR elements and promoters. In the looping model, availability, abundance, accessibility to LCR elements, and synergistic interactions of different transcription factors could create many different types of molecular machineries with different molecular functions. These machineries are easily fine-tuned by subtle changes in the LCR elements or the four features of transcription factors mentioned above. However, the linking model would require the uniformity of the protein complex linking the LCR elements and the promoters. It would

lay much more burden on B cells to express many different kinds of protein units that could exert different functions.

Although the available evidence seems to favor the looping model, the looping model should be modified to account for all of the experimental data. For example, the effect of promoter competition may depend on the direction of the competition. When the neo^r gene was inserted anyplace upstream of hs3b, whether to replace a dispensable cis-element or not, sterile transcription, along with class switch recombination, was abolished. (Cogne et al., 1994; Manis 1998). In a cell constitutively expressing a γ2a gene, expression of this IgH gene ceased when neo^r was inserted in place of hs1,2 (Lieberson et al., 1995). But when the neo^r gene was placed 2kb 3' to hs4, there was no detectable defect in the IgH locus *in vivo* (Manis et al., 2003).

1.3.4 The IgH locus is opened in a stepwise fashion

It has been suggested that the IgH locus is gradually opened during development to accommodate different recombination, hypermutation and transcription events at different cell stages (Chowdhury and Sen 2001).

The opening of the 5' locus begins at the region encompassing D_H through Cμ in prior to proB cells, followed by the opening of the 5'V_HJ558 family (about 2MB upstream of Cμ) induced by IL-7 stimulation, and opening of the D_H proximal V_H genes after DJ joining (Chowdhury and Sen, 2001). The 3'end of the IgH locus is opened at hs4~7 measured by histone H4 hyperacetylation in pro-B cells. In pre-B cells, this region is hyperacetylated in terms of hyperacetylation at both H3 and H4 at levels similar to Eμ. In splenic B cells, the whole hyperactylation domain extends 5' to cover the whole 3'IgH region, suggesting the more and more important role played by the 3'IgH regulatory

region in the latter stages of B cell development (Garrett et al., 2005). These results also show that the 3'end of the IgH locus, like the 5'end, is opened in a stepwise manner.

Overall, the chromatin in the IgH locus is gradually opened to prepare for the activation of the genes, and there also appear to be more than two chromatin states (open and closed) and the early states prepare for the later states. During the course of B cell development, B cells receive signals at various stages, and prepare the IgH locus in a stepwise manner for the transcriptional state that will culminate in maximum expression in plasma cells.

1.4 Significance

1.4.1 IgH locus as a model system for studying transcription regulation

The regulation of tissue-specific genes is one of many unique features of multicellular eukaryotic systems. It can be imagined that some of the molecular mechanisms of such regulation evolved from that of non-tissue specific genes in unicellular organisms, such as yeast, while others, like stepwise activation and maintenance are unique for tissue and stage-specific gene regulation.

Eμ is the first eukaryotic enhancer ever found. However, as a multi-gene model system, the LCR of the β-globin-like genes has been studied in more detail, probably because the β-globin locus is smaller and simpler than the IgH locus. In addition, the development of erythroid lineage is also simpler. Nevertheless, this complexity prompted us to further dissect the regulation of the IgH locus. It can be predicted that the discovery of novel mechanisms involve regulation of transcription distinct from those of the β-globin locus.

1.4.2 The utility in studying the mechanisms of other events that take place in the IgH locus, such as class switch recombination.

Following activation of surface Ig positive B cells, transcripts initiate from multiple sites upstream of the switch region (S region) of certain C_H constant regions. These transcripts run through the S region and terminate at the normal termination site of the C_H exons. An I exon is juxtaposed to downstream C_H exons during RNA processing. The I exon has multiple stop codons in all three reading frames. Thus, it is speculated that these transcripts are not translated into peptides and, therefore, are called germline transcripts or sterile transcripts. Accordingly, the process is called germline transcription or sterile transcription. Further analyses showed that germline transcription from certain C_H constant regions always accompanied switch recombination to those C_H regions following B cell activation (Lutzker and Alt, 1988). It was also shown that germline transcription and subsequent RNA splicing are required for class switch recombination (Jung et al., 1993; Lorenz et al., 1995). However, mechanisms underlying such requirement are currently unknown. Elucidation of these mechanisms would require manipulation of DNA sequences of I exon and, especially, the donor sequences for splicing. Such manipulations would be much more easily accomplished in BACs than in the endogenous locus.

The BAC is particularly useful in studying class switch recombination because the ability of a BAC to hold large pieces of DNA that encompass several constant regions makes BAC an ideal carrier of substrates for class switch recombination.

In studying the need for the of 3'IgH enhancers in class switch recombination, a previous study has shown that the enhancer pair hs3b and hs4 is required for class switch

recombination to most classes except IgG1 (Pinaud et al., 2001). A next logical step is to conduct single deletion studies of hs3b and hs4 to ask which one of these two enhancers is required for class switch recombination. However, the highly repetitive feature of sequences between hs3a and hs3b, and the strange sequence between hs3b and hs4 have precluded knocking out hs3b from its endogenous locus. In contrast, this has been easily achieved using simple homologous recombination in BAC.

1.4.3 Applications in studying stage-specific genes in the B lineage to gene therapy

The ultimate goal of the study of transcription of the IgH locus is to completely understand the mechanisms by which transcription is regulated both from V_H and I exon promoters in the locus at every stage of B cell development. These mechanisms include the following aspects:

a. Availability of transcription factors that can bind IgH regulatory sequences; Their combinatorial binding and functional interactions as well as their post-translational states.
b. The influence of chromatin structures and chromatin domains, including DNA and histone modifications.
c. Subnuclear localization of the IgH locus.

Each is a complex issue by itself. What make it more complex are the various kinds of interplay within and among these three considerations. The first step toward this goal is to define the control elements that are sufficient and/or indispensable for transcription at a particular cell stage. Achieving the first step itself has major scientific and biomedical importance.

First, if one could define the combination of control elements that become active only after a certain stage, it is possible to design tissue and stage-specific constructs that express the CRE enzyme to achieve conditional knockout in B lineage cells to study the roles of proteins that not only function at that particular stage but also are indispensable for B cell development at that stage. Examples of these proteins include Oct2, BSAP, B-limp, etc. Secondly, it is also of medical interest to be able to reconstruct the LCR so that people make use of it in gene therapy. For example, a virus can be constructed which contains a gene under the control of the IgH LCR so that this gene is specifically expressed in B cells at particular developmental stages in order to cure diseases resulting from genetic deficiency, e.g. CD40, exclusively expressed in surface Ig-positive B cells, can be used to treat X-linked hyper-IgM syndrome (HIMG1). Since the therapeutic gene is only ectopically expressed at a particular stage of B cell development, the therapy has fewer side effects.

An initial attempt has been made by Cogne's group to define the regulatory elements required for CSR (Pinaud et al., 2001). Although this study focused on DNA rearrangement (Class Switch Recombination), the same study showed that sterile transcription was also dependent on the pair hs3b & hs4. It would be interesting to see whether the pair of enhancers hs3b and hs4 are sufficient to drive transcription from an I exon promoter at the surface Ig positive stage. The focus of my project is plasma cells. In this study, I developed a transgene transcription system based on Bacterial Artificial Chromosomes. This enables us to easily modify the regulatory elements in an almost fully natural context. This will allow us to systematically study the regulatory elements *in vitro* and *in vivo*.

Chapter 2 Materials and Methods

2.1 Principle and procedures of BAC modification in bacteria.

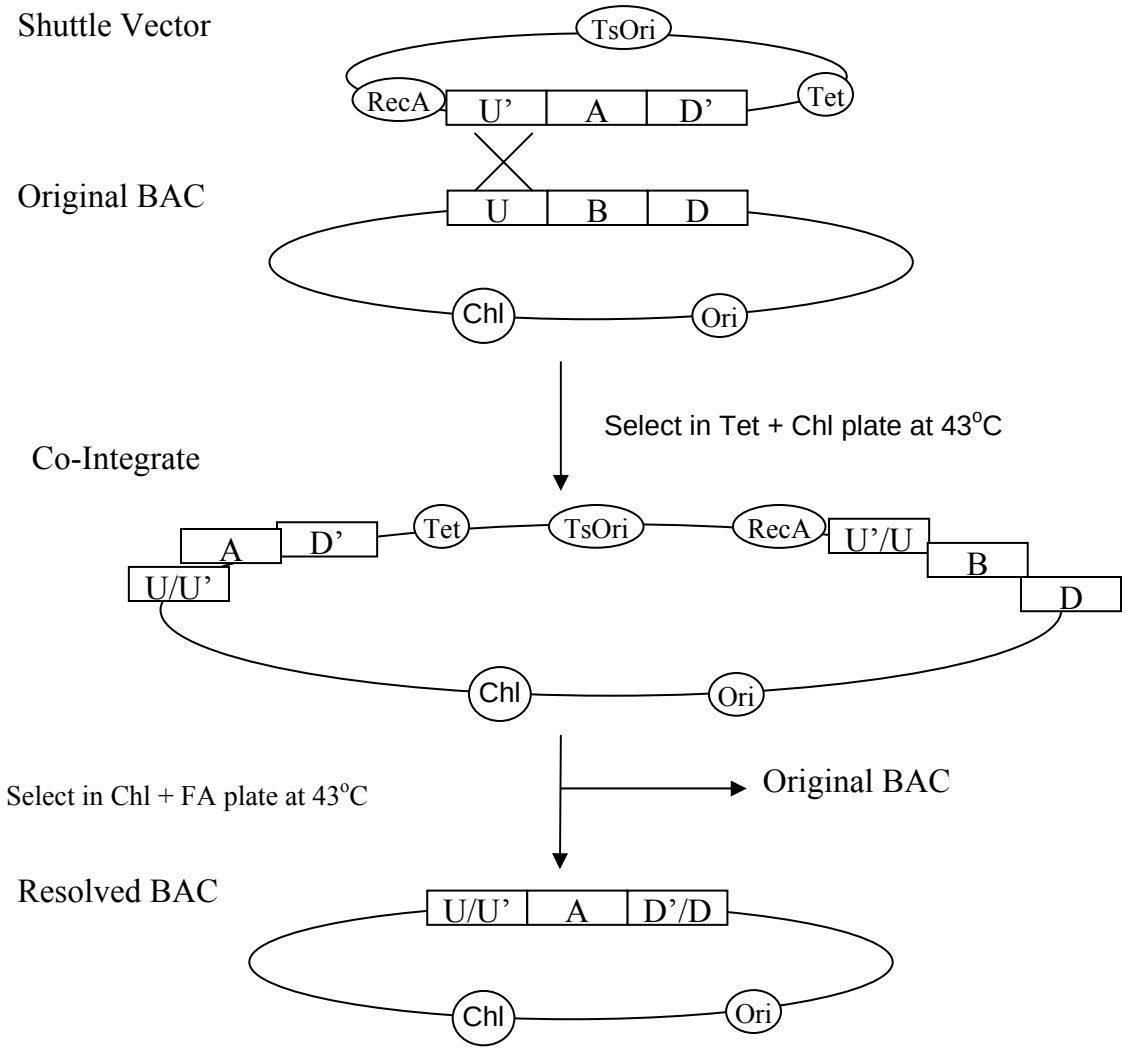
As shown in Figure 2-1, the goal of BAC modification is to replace fragment B in the original BAC with fragment A in the shuttle vector through homologous recombination of the U and D arms (upstream arm and downstream arm), each about 500bp. This modification cannot only accomplish replacement but also can carry out deletion (if A is nothing) or insertion (if B is nothing). First, the homologous arms U and D and the fragment A were cloned in the same orientation into a high copy plasmid called the building vector. The fragment containing U, A, and D were then cut out from the building vector and cloned into the shuttle vector which contains a temperature sensitive origin of replication which does not support replication at temperatures above 43^{o}C, a RecA gene, and a tetracycline resistance gene.

The recombinant shuttle vector was then transfected into competent cells containing the BAC to be modified. The host strain DH10B is of genotype recA^{-}Chl^{r} (chloramphenicol resistant). The lack of recA enzyme helps to prevent unwanted recombination and instability of the BAC DNA inside the bacteria. Taking advantage of the RecA protein provided by the shuttle vector, a first recombination event occurs, resulting in a co-integrated BAC carrying the whole shuttle vector. The co-integrated BAC contains both origins of replication and both tetracycline and chloramphenicol resistance genes. The origin of replication of the shuttle vector does not support replication at temperature above 43^{o}C. The original BAC does not confer resistance to tetracycline. So, when grown in an agar plate supplemented with both tetracycline and chloramphenicol at 43^{o}C, bacteria with only the unintegrated shuttle vector and original

BAC do not survive, while the co-integrant which can rely on the origin of replication from the original BAC and carries resistance genes to both drugs, can survive. Formation of the co-integrant can be confirmed by Southern blot probed with either of the two homologous arms, the U arm or the D arm. These two probes will give rise to a characteristic two-band pattern, rather than one which is shown in the original BAC. To accomplish the second round of homologous recombination, the co-integrants must lose the shuttle vector together with the B fragment through recombination with the arm not used in the first round of recombination. Co-integrated BAC is grown at 30^{o}C in chloramphenicol to allow this homologous recombination to occur. A population of bacteria with a second round of homologous recombination will have lost the shuttle vector, but contain only either resolved BAC or the original BAC. The colonies are spread on a plate supplemented with Chloramphenicol and fusaric acid at 43^{o}C. Since fusaric acid is toxic to those bacteria expressing the tetracycline resistant gene, those co-integrants which do not undergo the second round of recombination do not survive. Clones are picked and mini-preps of plasmid DNA are made. The DNA is subjected to Southern blot analysis to identify clones that contain the resolved BAC. If the homologous arms U and D are of relatively equal length, the outcome would be approximately half clones bearing resolved BAC and half bearing original BAC (Yang et al., 1997). Clones bearing resolved BAC are subjected to further modification or transfection experiments. An example of such modification is given in the first part of the Results chapter (Chapter 3): insertion of VDJ into the BAC 141e18 to form the BAC B1-8α.

Figure 2-1 BAC modification basic idea: The BAC modification is accomplished in two steps. First, a shuttle vector harboring fragment A flanked by two homologous arms U' and D' recombine with the original BAC, resulting in the co-integrated BAC with the whole shuttle vector inserted into the BAC. Since the origin of replication (TsOri) of the shuttle vector does not function at temperature above 43^{o}C, bacteria cannot contain independent shuttle vector not integrated into the BAC. In addition, the original BAC does not have the tetracycline resistance gene. Therefore, colonies that grow on double drug (Chloramphenicol and tetracycline) selection plates at 43^{o}C would only contain co-integrant. The selection of the second round of recombination involves the deletion of the shuttle vector from the co-integrant. Selection is based on the principle that tetracycline resistant bacteria are sensitive to Fusaric Asic. Recombination could occur between arms used in the first round of recombination (e.g. the U arm in this figure), resulting in the original BAC. Alternatively, it could also occur between arms not used in the first round of recombination (the D arm in this figure), resulting in the desired resolved BAC. Tet: Tetracycline resistance gene. Chl: Chloramphenicol resistance gene. FA: Fusaric Acid. TsOri: Temperature sensitive origin of replication that does not function at temperature above 43^{o}C.; Ori: Origin of replication.

Figure 2-1 BAC modification—Basic Idea



2.2Cell lines

9921 is an IgG2a-producing class-switch variant that was derived, through an intermediate, from the IgG2b-producing plasmacytoma MPC11 (Eckhardt and Birshtein, 1985). In the course of the heavy chain class switch, Eμ was deleted from the γ2a heavy chain transcription unit in 9921.

J558 is an IgA producing myeloma obtained from ATCC (ATCC, Item no. TIB-6)

9921 and J558 were maintained in Dulbecco's modified Eagle's medium (DMEM, Invitrogen 12100-061) with 10% Bovine Calf Serum (Gemini, CA, Catalogue no 100-506). The medium contained 50U/ml penicillin/streptomycin (Atlanta Biologic, catalog no. B21110) and 2mM L-Glutamine (Atlanta Biologicals, catalog no. B90310). Both cell lines were maintained at 37^{o}C in an atmosphere of 8% CO_2.

2.3 BAC, plasmid constructs and fragments:

In all plasmid descriptions, the total size of the described plasmid is provided in parentheses.

pSV1-RecA (11.0kb). Basic shuttle vector obtained from Dr. Nathaniel Heintz (The Rockefeller University, New York, NY) . It contains a temperature sensitive origin of replication which cannot initiate replication at temperatures above 43^{o}C, a RecA gene whose product catalyzes homologous recombination, and a tetracycline resistance gene (Yang et al., 1997). The homologous recombination unit is released from the building vector and cloned into the SalI site of pSV1 to form a working shuttle vector.

141e18 (92kb): a BAC that contains part of Cε (Starts from position 1430 of sequence X01857.1), the entire Cα gene and 80kb of sequence downstream of Cα (Ends at 91450 of sequence AF450245), which covers all the known IgH DNaseI hypersensitive sites and

four non-IgH genes (*hole*, *Crp*, *Crip2* and *Mta1*). This BAC was kindly provided by Dr. Siegfried Janz (NCI, NIH, Bethesda MD)

B1-8α (95kb): the prearranged VDJ transcription unit was inserted between IPAU and IPAD (see below) of 141e18 to generate a reporter BAC that is able to express IgA in B cells.

pBSIPAU (3.5kb): the upstream arm for homologous recombination in BAC IPAU was amplified and purified from BAC 141e18 using primer pair:

IPAUF: 5'- ACGCGTCGACCAGTAGGATGTGTAGAGGAT-3'

IPAUR 5'-CCGCTCGAGCCAGGACTCCACATGCAT-3'

The nucleotide sequences were obtained from Genbank accession number U08933. The two primers are designed to add SalI and XhoI sites to their ends, respectively (The restriction sites on the primers are underlined). Then IPAU was purified and cut with SalI and XhoI and cloned into the SalI site of the plasmid pBSIISK(+) (Stratagene Catalog no 212205) in the (-) orientation (the direction of the fragment is defined relative to the origin of replication of the plasmid pBSIISK)

pBSVDJ (+/-) (5.2kb): The rearranged VDJ fragment was cut out from the plasmid pIVHB1-8L2neo^r, a gift from Dr. Klaus Rajewsky (Harvard Medical School, Boston, MA) as a 2.2kb ClaI fragment and cloned into the ClaI site of pBSIISK(+). This VDJ gene segment was derived from the (4-hydroxy-3-nitrophenyl) acetyl (NP)-binding antibody B1-8. This V_HB1-8 was modified by a silent point mutation in codon 92, inactivating the "internal" heptameric recombination signal sequence (RSS) known to mediate the so-called V_H replacement reaction (Sonada et al., 1997). Two plasmids were obtained. One with (+) orientation—pBSVDJ (+) was used to construct the building

vector for B1-8α. The other in (-) orientation—pBSVDJ(-) was used to construct the building vector for BEμ.

pBSVDJIPAD (5.7kb): The downstream arm for homologous recombination in BAC, IPAD, was amplified and purified from BAC 141e18 using primer pair:

IPADF: 5'- CCGCTCGAGCTCAGTCTGACCCATCCACA-3'

IPADR: 5'-ACGCGTCGACAGCCACAACAGCCTGAGT-3'

The nucleotide sequences were obtained from Genbank accession number U08933. Similar to IPAU, IPADF and IPADR have XhoI and SalI sites at their ends, respectively such that IPAD has an XhoI site at its 5' end and a SalI site at its 3' end. Following amplification and gel purification IPAD was double digested with XhoI and SalI and cloned into the SalI site of pBSVDJ (+) such that IPAD lies 3' to VDJ and points in + orientation of replication of the plasmid.

pBSUVD (6.1kb): IPAU and VDJIPAD were cut out from their parental plasmids: pBSIPAU and pBSVDJIPAD, respectively, with XhoI and XbaI and cloned into the XhoI site of pBSIISK(+) through three-way ligation, generating the final building vector: pBSUVD with IPAU VDJ and IPAD pointing in the same orientation. This is the building vector for insertion of VDJ into 141e18, using homologous recombination.

pSVUVD (14.2kb): The three-component fragment UVD which contains IPAU, VDJ, and IPAD was digested with XhoI and cloned into the SalI site of pSV1RecA, generating the shuttle vector for insertion of VDJ into 141e18, using homologous recombination.

pBSCα (3.2kb): This plasmid contains the 300bp Cα fragment used as a probe to detect IgA messenger RNA in Northern blot. Cα was amplified with primer pair Cα Ch3F and

CαCH3R and cloned into the EcoRV site of pBSIISK (obtained from genbank accession no. D11468):

CαCh3F: 5'-GAATGAGCTCTTGTCCCTGAC-3'

CαCh3R: 5'-CATGATCACAGACACGCTGAC-3'

pCH3γ2a: This plasmid contains a 300bp SacI fragment covering the CH3 region of γ2a which is used as a probe to detect IgG2a messenger RNA in Northern blots (Shi and Eckhardt 2001).

pBSCα80 (3.4kb): A 500bp fragment 80kb downstream of Cα was amplified from BAC 141e18with primers :

Cα80F: 5'-GCATGTGCATTCAACAGTGG-3' (obtained from Genbank accession no. A450245)

Sp6: 5'-GCTATTTAGGTGACACTATAG-3'.

and cloned into the EcoRV site of pBSIISK(+).This probe was used to show whether the BAC is intact following transfection into cell lines.

pCε: This plasmid contains part of the third exon of ε constant gene and was used to show whether the 5' end of the BAC was intact following transfection into cell lines (Ishida et al., 1982)

pBSIPAD (3.5kb): pBSVDJIPAD was digested with ClaI and the VDJ fragment was released. The remaining part of the plasmid was self-ligated to form pBSIPAD.

pBSIPADΔC (3.5kb): In order to avoid too many repetitive sequences formed by the sequences lying in the multiple cloning site (MCS) of pBSIISK, the major part of the MCS was removed by double digesting pBSIPAD with SpeI and HindIII and then filled in with Klenow fragment of DNA polymerase to form blunt ends and religated.

pBSVDJEμ (6.2kb): A 1 XbaI fragment containing Eμ was cut out from puc18-Enh and blunt ended and cloned into the EcoRV site of pBSVDJ. Care was taken to make sure that Eμ was oriented in the same direction as VDJ.

pBSVEDΔC (6.2kb): VDJEμ was cut out from pBSVDJEμ with XhoI+XbaI+ BsaHI (BsaHI just to cut vector to smaller pieces for ease of fragment isolation). IPADΔC was cut out from pBSIPADΔC as an XhoI & XbaI fragment. The two fragments were ligated into the XhoI site of pBSIISK.

TopoVEDΔC (7.8kb): Because homologous recombination in bacteria requires that the upstream arm is similar in size to the downstream arm, a VEDΔC fragment with shorter upstream arm VDJ was amplified with primer pair using pBSVEDΔC as template:

VDJf: 5'-ACGCGTCGACTGAGCACACAGGACCTC-3'

IPADR: 5'- ACGCGTCGACAGCCACAACAGCCTGAGT-3'

The 2kb PCR product was cloned into TOPOBlunt4 vector using the Zero Blunt TOPO cloning kit (Invitrogen, catalog no. K2880-20) to generate TopoVEDΔC. This plasmid served as the building vector for insertion of Eμ into the BAC B1-8α.

pSVVEDΔC (14.7kb): Fragment VEDΔC was released from TOPOVEDΔC as an XhoI fragment and cloned into the basic shuttle vector pSV1RecA.

plxa and plxb (5kb): The usefulness of CRE-loxP mediated deletion has been exemplified by several knockout studies of cis-elements (Manis et al., 1998; Pinaud et al., 2001) as well as conditional knockouts. It is very useful if one could have a construct with loxP sites flanking a multiple cloning site. Therefore I started with a simple approach to build two simple constructs with loxP flanking restriction endonuclease sites: plxa and plxb. Plxb was used in construction of almost all important plasmids in this

project. The construction scheme was: loxP flanked neo^r was cut out from ploxP2neo^r (A gift from Dr. F.W.Alt from Havard Medical School, Boston, MA) with NotI and cloned into pBSIISK (+) cut with SpeI and ClaI and blunt ended. There is no need for neo^r in either plxa or plxb except that when neo^r is cut out with two enzymes, two restriction sites are generated in the vicinity of the two loxP sites. Neo^r could be cut out with BamHI and EcoRI and one can easily clone in any gene wanted between the two loxP sites. Distance between sites: NotI—loxP: 36bp; loxP—EcoRI: 187bp; BamHI—loxP:390bp; loxP—NotI: 32bp; loxP sites are 32bp by themselves.

2.3 Plasmids used to first remove hs4 from BAC B1-8α using homologous recombination.

The aim of the project was to remove hs4 from the transgene in its original integration site in transfectants. Therefore the construct should contain loxP flanked hs4. In order to flank the hs4 in the BAC with loxP sites, we decided to remove hs4 from the BAC B1-8α first and then insert back a loxP flanked hs4. The scheme of the construction of the building vector to delete HS4 is similar to that to insert VDJ into the BAC 141e18.

pBSHS4U (3.7kb): HS4U is a 750bp fragment directly downstream of hs3b. This 750bp was chosen because we and others have failed to clone or amplify the intact 3kb fragment between hs3b and hs4, probably due to the highly repetitive sequences within this region. hs4U was amplified from BAC B1-8α with primer pair (Genbank accession# AF450245):

HS3bF: 5'-ATGCAAGCTTAATGGATGTGAGATGAGG-3'

HS4UR: 5'-CATGTTGAACAGTGAGTCTGCAG-3'

Topo3'hs4-3kb (7kb): a 3kb fragment downstream of hs4 was amplified using primer pair and cloned into TOPO vector (Genbank accession# AF450245):

HS4DF: 5'-CCTACCCACCTAACTCCAAGC-3'

HS4DR: 5'-GGTAGGAAGGGCGAATTCGC-3'

pBSHS4D (5kb): A 2kb fragment HS4D cut with BamHI and HindIII from Topo3'hs4-3kb and inserted between the BamHI and HindIII sites of pBSIISK.

pHS4UD (5.8kb): HS4U and HS4D were cut out from pBSHS4U and pBSHS4D with both XhoI and XbaI, and cloned into the XhoI site of pBSIISK(+), using three-way ligation, generating the plasmid pHS4UD, in which HS4U and HS4D are aligned in the same orientation.

pSVHS4UD: The fragment that covers HS4U and HS4D was cut out from the plasmid pHS4UD with XhoI and cloned into the SalI site of the shuttle vector pSV1RecA, generating the shuttle vector pSVHS4UD, which was used to generate the BAC B1-8αΔhs4

2.4 Plasmids used for insertion of loxP flanked HS4 into the BAC B1-8αΔhs4 to generate the BAC Blxhs4.

Pcr2β^s: pcr2β^s was a gift from Dr. Karl Drlika (Public Health Research Institute, New York, NY). It contains the complete 626bp cDNA of the sickle cell anemia form of beta-globin (Gene Bank Accession no. V00497).

pEβ^s: a 420bp fragment containing nucleotide 1 through 418 was cut out from Pcr2β^s as an EcoRI fragment and cloned into the EcoRI site of pEGFP-C1 (Catalog no. 6084-1, Clontech, Palo Alto, CA). The purpose of cloning into this plasmid was to take advantage of the XhoI and SalI sites in pEGFP-C1 (see below).

plxhs4: a 1381bp HindIII-PstI HS4 fragment was cut out from the plasmid HS3,4 which was a gift from Dr. Mark Groudine (University of Washington, Seattle, WA) (Madisen,

1994) and cloned into plxb BamHI and EcoRI to generate plxHS4, with the HS4 pointing to the (-) orientation of the plasmid.

plxhs4b: β^s fragment was cut out from pEβ^s as an XhoI/SalI fragment and cloned into the SalI site of plxhs4. A clone was chosen such that the β^s was pointing at the (-) orientation of the plasmid, and that the SalI site at the 3' end of β^s was regenerated while the XhoI site which was at the 5' end of β^s was destroyed when it annealed with the SalI site of plxhs4.

TOPO-U4βD: the fragment u4βd was generated piecing together hs4U, lxhs4β and hs4D fragments using the PCR strategy described in (Misulovin et al., 2001) and cloned into TOPOBlunt4 vector. Briefly, the 500bp upstream arm HS4U fragment was amplified using pBSHS4U as template with primer pair

hs4u3f: 5'-CATCAGACTCGAGCAGTAGCCAGTTACCAGGAC-3'

hs4u3r: 5'-CCTAGTTCTAGACACCTGCCTCTGACTCCTGA-3'

lxhs4b fragment was amplified using plxhs4b as template with primers

lxhs4bf: 5'-CAGAGGCAGGTGTCTAGAACTAGGGCCGCTGG-3';

lxhs4br: 5'-TGAGGTCTACACCGAGGTCGACTGCAGAATTC-3'

The downstream arm HS4D was amplified using pBSHS4D as template with

HS4D3f 5'-CAGTCGACCTCGGTGTAGACCTCAAAGTCAGCC-3'

HS4D3r: 5' CATCAGACTCGAGGGTGTCTCTGTGTCTGTTC-3';

The primers were designed in such a way that after amplification, HS4U and lxhs4b share a 24bp DNA junction sequence at the 3' end of HS4U and 5'end of lxhs4b. lxhs4b and HS4D share 24bp DNA junction sequences at the 3'end of lxhs4b and 5'end of HS4D. All three fragments were purified using Qiaguick^R Gel Extraction kit (Catalog no. 28704,

Qiagen, CA) and were put together in a PCR reaction so that the two junction DNA sequences linked the three fragments together with end primers HS4U3f and HS4D3r to generate the fragment u4βd. HS4U3f and HS4D3r bear 5'XhoI recognition sites so that the fragment u4βd has XhoI sites at both of its ends. Care was taken to avoid the XhoI site at the end of plasmid plxhs4β. The u4bd fragment was then cloned in TOPOBlunt4 Vector from Invitrogen.

pSVu4βd: The u4βd fragment was released from TOPOu4βd by XhoI digestion and cloned into the shuttle vector PSV1-RecA to generate pSVu4βd for recombination with the BAC B1-8αΔhs4.

pEGFPCRE: In order to delete HS4 from the BAC transgene in transfectants more effectively, we decided to clone a plasmid that expressed both EGFP and CRE together.

plasmid carrying EGFP and CRE enzyme.

pBS185: is a CRE recombinase gene expression vector (Gibco BRL, Grand Island, NY, catalog no. 10347-011) in which CRE is expressed under the control of the cytomegalovirus (CMV) promoter. CRE recombinase mediates loxP site-specific recombination.

pEGFP-C1: encodes a red-shifted variant of green fluorescence protein (catalog no. 6084-1, Clontech, Palo Alto, CA).

pEGFPCRE (8.8kb): a 3.4kb fragment of CRE was cut out from pBS185 with HindIII and cloned into the HindIII site of pEGFP-C1 in the same orientation as EGFP.

2.5 Stable Transfection of BACs into the 9921 cell line

All the BACs transfected were digested with NotI. The BAC inserts (with very large sizes) were purified from the 6.7 BAC vector using gel filtration. Briefly, a gel filtration column was made by loading 12~13ml sepharose CL-4B beads (Catalog no. 17-0150-01, Amersham Pharmacia Biotech, Uppsala, Sweden) into Platinum-Cured Silicone Tubing (catalog no. 420400-0730, Nalgene, Rochester, NY). The column was washed with ddH_2O for three hours before sample BAC with bromophenolblue dye was loaded. After the sample completely entered the column, fractions of 0.5ml were collected. Usually 12 fractions were collected, and fractions with intact linearized BAC were pooled and concentrated by speedvac for transfection into 9921 using electroporation. 10μg BAC and 400ng pBSSKneo^r (Shi and Eckhardt 2001) were mixed with a 1ml suspension of 10^7 9921 cells and transferred into a 0.4cm (width) electroporation cuvette (Biorad, Hercules, CA). An electric pulse was delivered at 960μF and 250V by a Bio-Rad Gene Pulser^{TM} electroporator and capacitance Extender TM (Bio-Rad, Hercules, CA). The cells were then diluted in non-selective medium and plated at various densities into 96-well tissue culture plates. The most often used density was 8000 cells/well. 48 hours following the transfection, medium supplemented with G418 (catalog no. 11811-031, GibcoBRL, Grand Island, NY) was used to replace half of the medium, so that the concentration of G418 in the selective medium reached 1.5mg/ml in order to select for G418 resistant clones. Colonies were visible about 2 weeks after transfection. Transformants arose in about 30~40% of the individual wells on each culture plate, indicating that most growing wells contain progeny of a single transformant. Subcloning

was conducted for cell lines transfected with B1-8α, BEμ and Blxhs4 to guarantee each line analyzed was of a pure population.

After the clones were established, the BACs were subjected to PCR and Southern analysis to determine whether they were intact. The 5' ends of BACs were amplified by

T7: 5'-GTAATACGACTCACTATAGGG-3'

CepsilonR: 5'-GGACGACATGACTTAACCAG -3' (accession number X01857.1)

The 3' ends of the BAC were amplified by:

Sp6: 5'-GCTATTTAGGTGACACTATAG-3'

Cα80F: 5'-GCATGTGCATTCAACAGTGG-3' (Genbank accession no. A450245)

Following CRE-loxP mediated deletion, the region surrounding hs4 was amplified by:

hs4uf: 5'-CTAGCCAGGCAGTGATAGC -3'

β-globinr: 5'-TGAGGTTGCTAGTGAACACAG-3'

The 700bp fragment was gel purified and subject to sequencing analysis to confirm that the deletion was due to CRE-mediated loxP recombination.

Southern Blot Analysis

Restriction enzyme-digested DNA was loaded into each lane of a 0.7% agarose gel. For selection and analysis of BAC DNA, 0.5~1μg of digested DNA was used. For genomic DNA, ~25μg of digested DNA was used. Size-fractioned DNA was then transferred to nylon membrane (Cat no. N00HY00010, Micron Separations Inc, Westborough, MA) or (Cat no. 10415296, Schleicher and Schuell, Bioscience). Blots were hybridized either with traditional methods or Quickhyb (Cat # 201220, Stratagene, CA). For traditional methods, blots were pre-hybridized and hybridized at 63^{o}C in buffer

containing 7.5X Denhardt's solution, 3X SSC, 100μg sonicated salmon sperm DNA and 0.5% SDS. Probes were labeled with random primer labeling methods with MegaPrime™ labeling kit (cat no. RPN 1605, Amersham, Arlingon Heights, IL). To remove non-specifically bound probe, blots were washed twice in 1X SSC and 0.1% SDS at 65°C for 30 minutes. The probes used are indicated in each individual figure.

PCR.

Polymerase chain reactions (PCR) were conducted with either HotStart Taq polymerase (Catalog no. 203203, Qiagen, CA) (for detection) or Pfu Turbo (Catalog no. 201220, Stratagene, CA)(for cloning and sequencing). Most amplifications were done in 30 cycles using GeneAmp PCR system 9600 (Perkin Elmer, MA).

Northern blot analysis

Total cellular RNA was isolated by Trizol reagent (cat no. 15596-026, GibcoBRL, Grand Island, NY) according to manufacturer's instruction. Approximately 20μg total RNA was analyzed. RNA was denatured with formamide and size-fractionated by electrophoresis through 1% formaldehyde-agarose gels. Following electrophoresis, RNA was transferred to Genescreen nylon membrane (cat no. 881623-1200, PerkinElmer, MA). Blots were pre-hybridized with pre-hyb solution for two hours and hybridized to ^{32}P-labeled DNA probes at 42°C overnight in a pre-hyb solution of 50% formamide, 2.5X Denhardt's solution, 5X SSC, 0.1% SDS, 50mM $NaPO_4$ (PH7.4), 50μg/ml poly[A], and 50μg/ml sonicated salmon sperm DNA. To remove non-specifically bound probe, blots were washed in 1XSSC, 0.1% SDS at 65°C twice for 30 minutes.

Enzyme-linked immunosorbent assays (ELISA).

Microtitre plates (Dynatech Laboratories, Chantilly, VA) were coated with 2μg/ml polyclonal Rabbit anti-mouse IgA (cat no. 61-6700, Zymed, CA) or Goat anti Mouse IgG (cat no. 62-6600 Zymed, CA). Coated wells were incubated with 100μl samples of diluted culture supernatant (IgA 1:100 dilution, IgG 1:400 dilution). The presence of expressed proteins was detected by enzyme-linked detecting antibodies Horseradish Peroxisdase (HRP) conjugated Goat anti mouse IgA(cat no. 62-6720 Zymed,CA) or HRP-Goat x mouse IgG (cat no. 31430 Pierce, IL) and subjected to chemiluminescent reaction with substrate ABTS single solution (cat no. 00-2024, Zymed, CA). The plates were then read at wavelength 405nm using Power wave 200 microplate scanning spectrophotometer (Biotek, MA) and analyzed by the software KC4.

Chapter 3 Results

3.1 Overview of Experiments:

To help the readers understand the whole of the experiments, an overview of experiments and a streamlined description of the BAC modifications are shown in Figure 3-1 and Figure 3-2. The large BAC insert consisting of a transgene that encodes an α heavy chain was transfected into the plasmacytoma cell line 9921. The number of copies of the transgene in each cell clone was obtained by Southern blots followed by phosphoimager analysis. Transcription levels of the transgene (α heavy chain) and endogenous IgH gene (γ2a heavy chain) were measured by northern blot followed by phosphoimager analysis. 9921 is a plasmacytoma cell line in which there are three untranscribed and translocated IgH loci that are juxtaposed with the *c-myc* gene and one copy of a productive rearranged IgH locus that gives rise to a γ2a transcript and secretable IgG2a immunoglobulin protein. The transcribed allele has lost its Eμ during class switch recombination (Eckhardt and Birshtein 1985). Therefore this γ2a gene is under the control of the 3'IgH regulatory region, serving as a good internal control for gene expression in the same cell line (Overview of experimental strategy in Figure 3-1).

Figure 3-1. Experimental system. We took advantage of the fact that the only IgH expressing allele of the plasmacytoma cell line lacks Eμ, making it comparable to the BAC constructs. Original BACs were modified to produce desired constructs—resolved BACs. Resolved BACs can serve as "original" BACs and can be subjected to further modification. The desired BAC constructs were transfected into 9921. Transcription level of IgA—the heavy chain expressed by the BAC transgenes—was measured and compared to that of IgG2a, which is expressed by the endogenous locus in 9921. Blxhs4 transgene clones were transiently transfected with GFPCRE plasmid to undergo CRE-loxP mediated deletion, so that the expression levels before and after hs4 deletion could be compared.

Figure 3-1: Overview of Experimental System

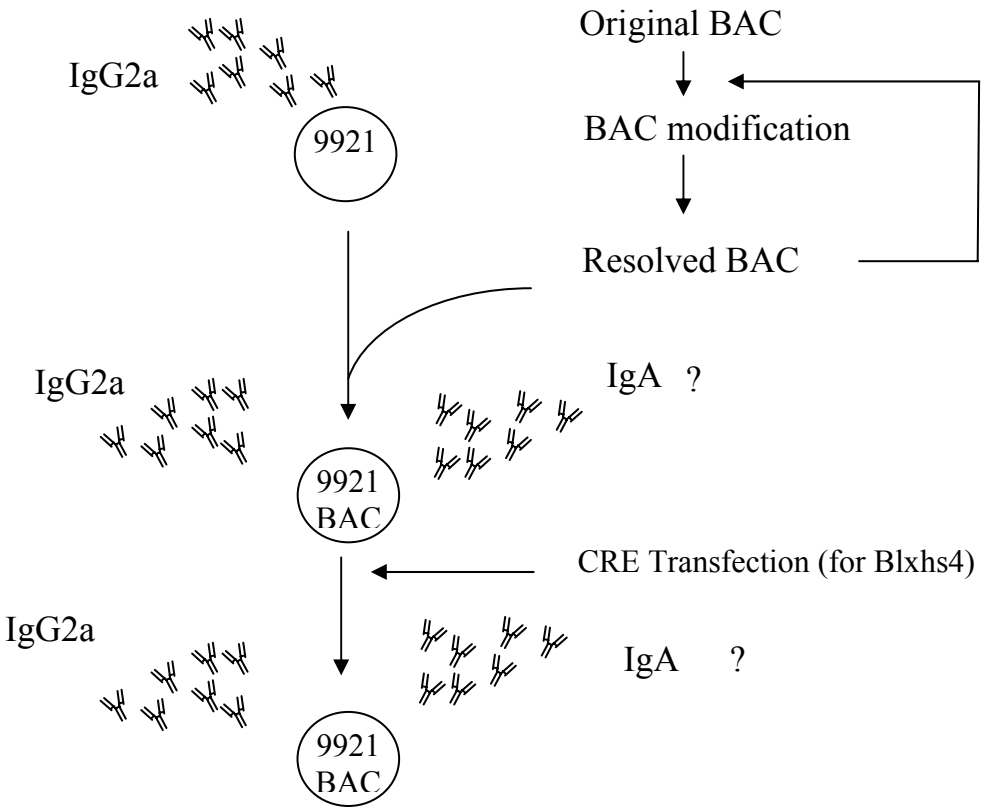
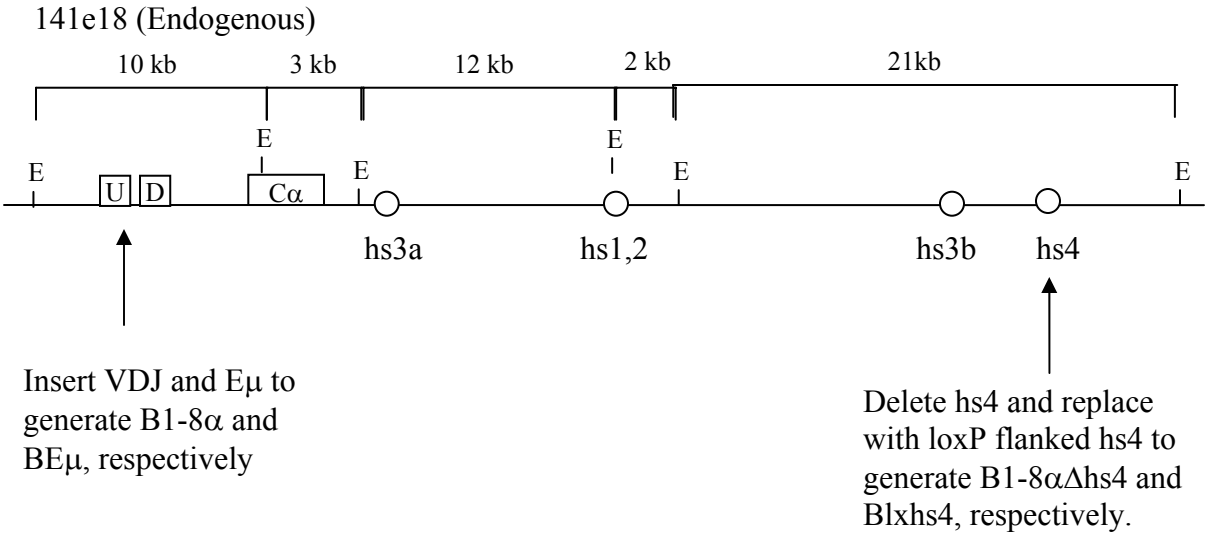
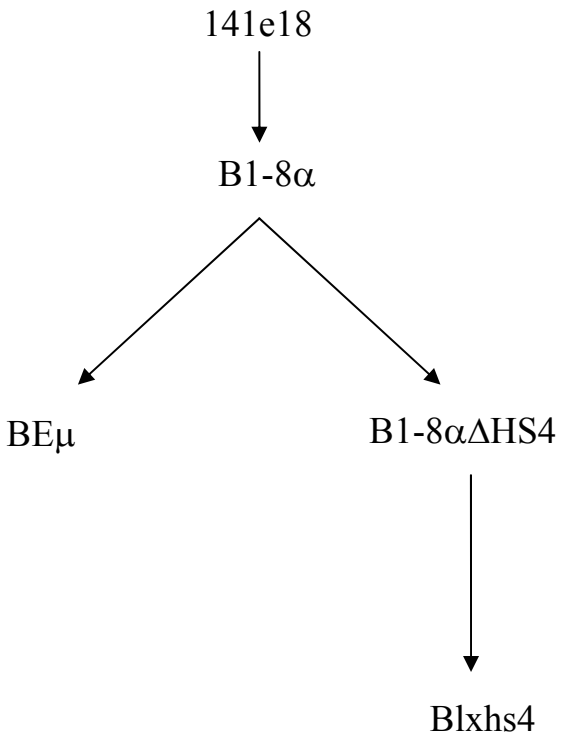


Figure 3-2. Streamlined view of BAC modifications. 141e18 is a BAC we received from Dr. Siegfried Janz from the National Cancer Institute, NIH. A rearranged VDJ was inserted upstream of Cα in 141e18 to form the BAC B1-8α. Eμ was inserted between VDJ and Cα to form the BAC BEμ. hs4 was deleted from B1-8α to generate B1-8αΔHS4. Floxed hs4 (loxP flanked hs4), followed by a DNA tag consisting of a portion of the human β-globin cDNA sequence, was inserted into B1-8αΔHS4 to form the BAC Blxhs4. For detailed restriction maps of these maps please refer to figures 3-3, 3-9 and 3-15.

Figure 3-2 streamlined view of BAC modifications





As shown in Figure 3-2, we started with the BAC 141e18, and generated four modified BACs by homologous recombination. They are: B1-8α, B1-8αΔhs4, Blxhs4, BEμ. All modifications were achieved using homologous recombination in bacteria. A rearranged VDJ fragment corresponding to the B1-8 variable region was inserted into 141e18 to generate B1-8α. hs4 was deleted from B1-8α to generate B1-8αΔhs4. A loxP flanked hs4 fragment was inserted into B1-8αΔhs4 to generate Blxhs4. Eμ was inserted 3' to the rearranged VDJ of B1-8α to generate BEμ. The homologous recombination events are described in a later part of the Results. The details of plasmid constructions are described in Material and Methods.

3.2 Generate a reporter BAC—B1-8α

Studies on the contributions, requirements, as well as mechanisms of transcription control by cis-elements in the IgH locus have been conducted at several levels, using a variety of approaches. Transient transfection assays conducted using individual IgH enhancers led to the discovery of the 3'IgH enhancers and the determination of the sequences of each enhancer (Pettersson et al., 1990; Lieberson et al., 1991; Matthias and Baltimore 1993; Michaelson et al., 1995; Madisen and Groudine 1994). Transient transfection assays with combinations of the enhancers suggested the existence and the power of synergy among the 3'IgH enhancers (Madisen and Groudine 1994, Ong 1998). EMSA assays on individual or subfragments of individual enhancers revealed the contributions of, the requirement for, and the interplay among numerous transcription factors (Michaelson et al., 1995; Singh et al., 1996; Sepulveda et al., 2001). Stable transfection assays using conventional plasmids showed synergy among the 3'IgH enhancers in the context of chromatin and suggested the possible LCR activities of the 3'

IgH region in cell lines representative of certain stages of B cell development (Madisen and Groudine, 1994; Shi and Eckhardt., 2001). Transgenic mouse studies showed that the 3'IgH enhancers were able to open the locus in the course of *in vivo* B cell development (Arulampalam et al., 1994b; Chauveau et al, 1999). Systematic knockout studies revealed requirements in class switch recombination (CSR) in the IgH locus (Cogne et al., 1994; Manis et al., 1998; Pinaud et al., 2001)

While all of the above approaches have provided a large amount of information and have helped us to understand some of the mechanisms underlying transcriptional control in the IgH locus, each has its own limitations. Transgenes constructed in conventional plasmids are often under the influence of neighboring chromatin structures and of control elements such as enhancers and silencers of adjacent genes. As a result, expression analysis on control elements using these plasmid expression systems may be distorted by the influence of the environment surrounding different transgene integration sites. In addition, factors other than DNA sequences may also be distorted when manipulating the constructs. For example, there are inverted repeats surrounding hs1,2. The possible secondary structure formed by these inverted repeats instead of the primary nucleotide sequence may play a role in regulating chromatin structure as well as transcriptional activity so that plasmids containing only the hs1,2 enhancer itself without the flanking repeats might not accurately measure this enhancer's activity. Furthermore, it cannot be ruled out that there exist uncovered control elements located between the enhancers, making it difficult to interpret the results of enhancer knockout studies. On the other hand, the dispersion of so many control elements over a long expanse of DNA and the redundancy among these control elements calls for manipulation of these elements at

multiple places in the locus. While knockout studies within the genome of mice provide genuine information that furthers our understanding of the requirements for these elements in controlling transcription in the IgH locus, these studies involve a rather time-consuming procedure that occludes us from conducting systematic studies by making changes in multiple places.

Therefore, we took advantage of the bacterial artificial chromosome (BAC) in order to develop a better approach for studying the IgH locus. BAC is one of the artificial chromosomes used in physical mapping and manipulation of genes with regulatory regions of large sizes (Monaco and Larin, 1994). It is based on the *E.coli* Fertility plasmid and maintained at one copy in recombination deficient bacteria (O'Connor et al., 1989). It has several advantages over the Yeast Artificial Chromosome (YAC): high cloning efficiency, high stability, and ease of both isolation and modification. An homologous recombination method was developed to modify BACs and subject BACs to transgenic study (Yang et al., 1997). It was shown that expression of genes regulated by sequences that span more than 20kb cannot be achieved using conventional plasmid vectors but can be achieved using BACs, which hold up to 300kb DNA (Yang et al., 1997; Yu et al., 1999). In addition, BACs provide a tool to manipulate cis-regulatory elements in functional studies, which is demonstrated in the present study.

The 5' end of the BAC 141e18 was mapped to the second exon of Cε and the 3' end was mapped to a site 80kb downstream of Cα. In order to build a testable expression system, we inserted an assembled VDJ (V_H gene) upstream of Cα through homologous recombination in bacteria (methods described in Yang et al., 1997). Briefly, upstream and downstream arms for homologous recombination IPAU and IPAD were amplified,

respectively, and were cloned with VDJ between them into a shuttle vector, generating the shuttle vector pSVUVD (Please refer to Material and Methods for details of plasmid construction). Recombinations in bacteria were carried out to insert VDJ between the two arms in the BAC. The two homologous arms, IPAU and IPAD, were by design, flanking the promoter for α sterile transcripts, IPα. As a result, IPα is deleted following BAC modification so that transcription rate is only observed in the context of the V_H promoter and the 3'IgH regulatory region. Please refer to figure 2-1 for principles of BAC modification: The goal is to replace the IPα (fragment B in figure 2-1) with a rearranged VDJ (fragment A in figure 2-1). The restriction maps of the original BAC-141e18, two possible co-integrants, and resolved BAC are shown in Figure 3-3. First, the shuttle vector is integrated into the original BAC to form cointegrants. The correctly formed co-integrant is isolated under the selective pressure of high temperature (43^oC) and two drug resistance markers: Tet^r and Chl^r (Please refer to Material and Methods for the principles of the selection by temperature and drugs). As shown in Figure 3-3, BamHI digestion and hybridization with IPAU should generate a 14kb fragment in BAC 141e18, but a combination of 11kb and 4kb fragments for one kind of cointegrant and a combination of 13.7kb and 0.5kb fragments for the other. Restriction analysis of the co-integrant used in these studies is shown in Figure 3-4. The second step is to delete the shuttle vector backbone along with either fragment A (VDJ) or fragment B (IPα). Accordingly, the resultant BAC could be the resolved BAC (if fragment B is deleted) or the original BAC (if fragment A is deleted) (Figure 2-1). The resolved BAC is selected based on Southern blot analysis. As shown in figure 3-3 and figure 3-5, after the recombination, a 14kb BamHI fragment that hybridizes with IPAU in 141e18 should become 4kb in the resolved

BAC. The resolved BAC was termed B1-8α (Figure 3-6). Because a BAC is a very large piece of DNA, it is possible to introduce unexpected changes in DNA sequences during homologous recombination in bacteria. To rule out this possibility, the whole BAC was digested with EcoRI, fragments were size-fractionated on an agarose gel and stained with ethidium bromide. There was no change in restriction pattern apart from the predicted changes (Figure 3-3 and Figure 3-6). The changes are as expected: a 10kb band from the original BAC 141e18 disappeared, a 3.5 kb band containing IPAU and an 8kb band containing IPAD from the resulting BAC B1-8α appeared.

After making sure that B1-8α had only the desired changes in its DNA sequences, the BAC was digested with NotI and the insert was purified (92kb for B1-8α) through a CL-4B column before it was transfected into the plasmacytoma cell line 9921(Methods described in Yang et al., 1997). Neo^r clones were screened for the presence of intact BAC by PCR and Southern blot. To determine copy number, genomic DNA from each clone was digested with HindIII, size-fractionated and subjected to Southern blot hybridization with a probe consisting of IPAU—the same way that the copy number of BAC BEμ was determined (Next section). As shown in Figure 3-12, the endogenous locus generated a 4.4kb band while the transgene generated a 3.7kb band. Intensities of bands were measured, using a phosphoimager and were used to calculate copy number. Since there are four chromosomes in 9921 that carry the 4.4kb HindIII restriction fragment detected with IPAU, transformants with a single copy of the BAC were expected to show hybridization of the 4.4kb:3.7kb bands at a 4:1 ratio. The transcription levels of α and γ2a heavy chains of each clone were measured by northern blot followed by phosphoimager analysis. In an initial experiment, B1-8α was

transfected into 9921 cells, and five stable transformants were recovered. As shown in Figure 3-7A, individual clones expressed varying amounts of BAC-encoded α–heavy chain mRNA. As described in a later section, we subsequently generated a modified version of B1-8α in which the 3'IgH enhancer hs4 was surrounded by loxP sites. 9921 cells were also transfected with this BAC (Blxhs4) and stable transformants isolated. As shown in Figure 3-7B, again, individual clones produced varying amounts of BAC-encoded α-chain mRNA. As there was no apparent change in the behavior of the BAC, before and after the modification to insert loxP sites, detailed quantification was done on the Blxhs4 transformants. The amounts of α mRNA and γ2a mRNA were separately normalized to the internal control GAPDH. The per copy normalized levels of γ2a and α heavy chain in these transformants are shown in Figures 3-8A and 3-8B, respectively. The amount of α mRNA and γ2a mRNA were separately normalized to the internal control GAPDH (Figure 3-7). According to these two plots, the variation of α chain is much greater than the γ2a heavy chain which is expressed endogenously. The relative expression levels in Figure 3-8C were taken as the ratio of α heavy chain over γ2a heavy chain in order to control for the differences in γ2a mRNA levels in different transfected clones. This ratio divided by copy number shows the per copy expression level of the transgenes. As shown in Figure 3-8C, the expression levels of α chain mRNA from one copy of BAC Blxhs4 ranged from 0.7% to 60% of the endogenous γ2a mRNA. Most of the clones produced α mRNA per copy at 20% γ2a heavy chain mRNA levels. A closer look revealed that the Blxhs4 transfectants can be divided into three distinct groups according to their per copy number expression levels: A low expressing group that expressed lower than 3%; An intermediate expressing group that expressed about

15%~20% and a high level expressing group that expressed about 60% of the endogenous level ($p<0.0005$). This result suggests that there are three kinds of integration sites that would permit the transgenes to be expressed at three different levels. A dot blot showing the expression level vs copy number is in Figure 3-8D, and t-test shows that the correlation between expression level and copy number is not statistically significant ($p>0.1$). Therefore, the BACs B1-8α and Blxhs4 do not behave so differently from the mini-cassettes constructed before (Shi and Eckhardt., 2001) with respect to transcription enhancement. As in that early study, the BAC Blxhs4 was expressed (per copy) in most clones at below 30% of the endogenous gene, expression level varied substantially from clone to clone, and the transgene expression level did not linearly correlate with copy number. The intact 3'IgH regulatory region in its natural configuration, therefore, does not confer transgene copy number dependent expression, and does not behave like a full LCR.

Figure 3-3. Restriction maps of BAC modification for integrating VDJ fragment between IPAU (U) and IPAD (D) of original BAC 141e18 to generate resolved BAC B1-8α. Both possible forms of the co-integrants are shown in the middle. Hatched boxes indicate arms (U and D) used in homologous recombination. U/U', U'/U, D/D', and D'/D indicate that the DNA fragment is a hybrid between BAC sequence and shuttle vector sequence following recombination. Sizes of predicted restriction fragments are indicated in brackets with probe used to detect the fragment in parentheses (U: IPAU; D=IPAD). Fragments containing the coding sequences for VDJ and RecA are shown in open boxes. Ovals represent the 200bp DNA sequence, corresponding to 4037 to 5037 of sequence U08933, encompassing the promoter of Cα sterile transcripts (Ip) which is deleted during insertion of the 2.2kb fragment containing VDJ coding sequences. As a result of this modification, the assembled VDJ gene segments lies 5.8kb upstream of Cα. Ts: temperature sensitive origin. Chl^r: chlroamphenicol resistance gene; Tet^r: tetracycline resistance gene; Restriction enzymes: B, BamHI; E, EcoRI; N, NotI.

Figure 3-3. Restriction maps of BAC modification for integrating VDJ fragment

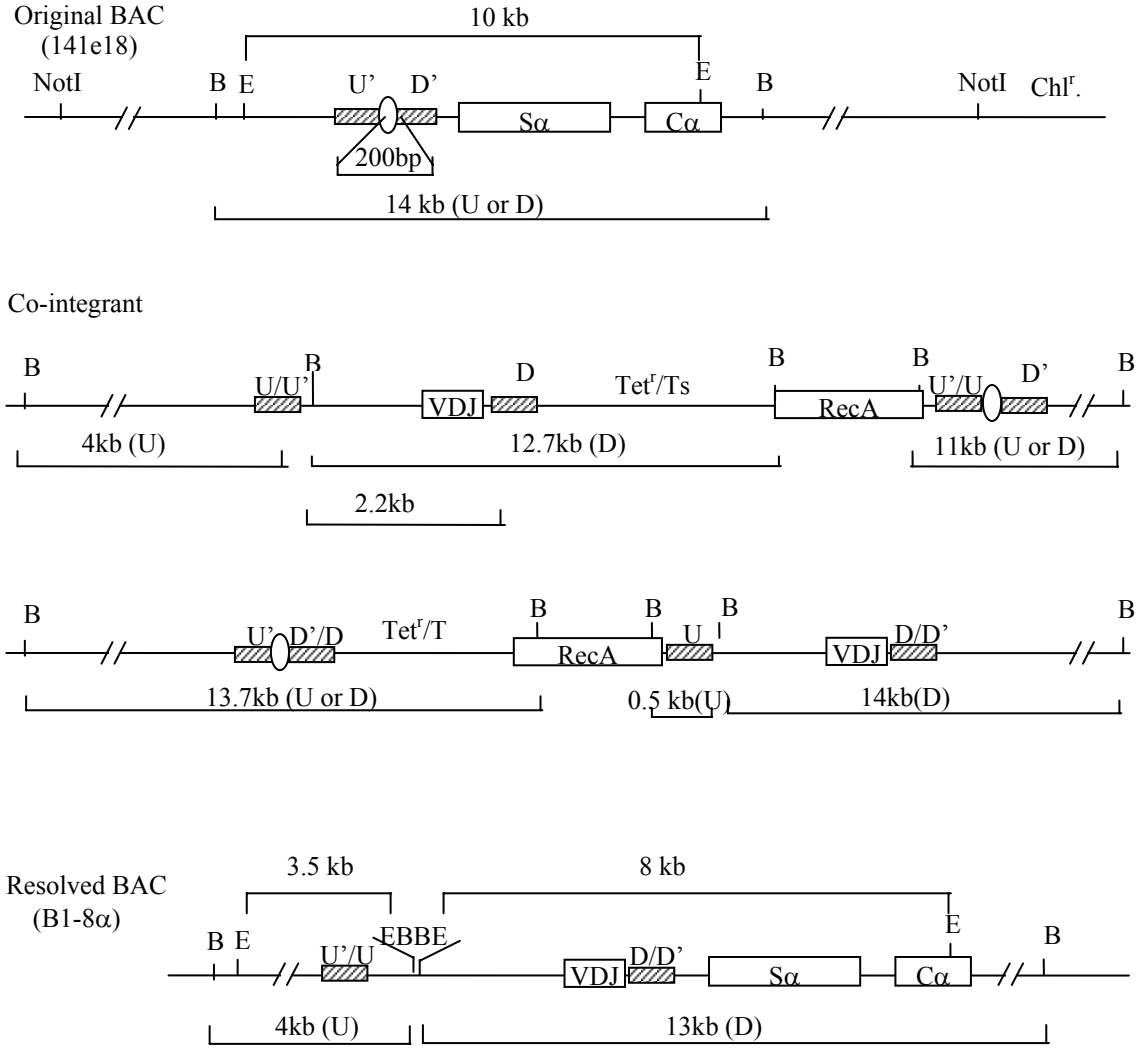


Figure 3-4. Identify co-integrants. BAC clones were digested with BamHI and hybridized with IPAU. The original BAC 141e18 generated a 14kb band while the co-integrant generated two bands: 11kb and 4kb. This revealed that in this particular clone, the "U" arm of the shuttle vector recombined with the "U" arm of the BAC in this first round of recombination.

Figure 3-4. Identify co-integrants

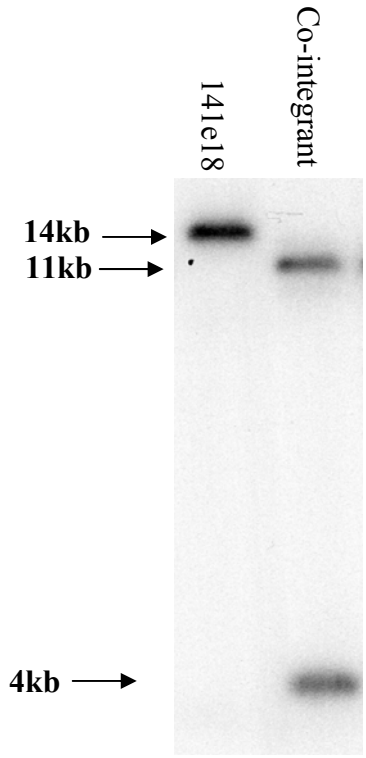
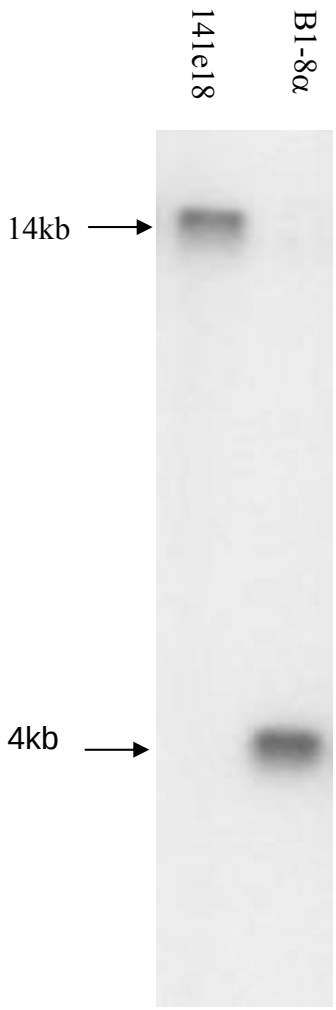


Figure 3-5 Identify resolved BAC. BACs were digested with BamHI and hybridized with IPAU. The original BAC 141e18 generated a single 14kb. The resolved BAC B1-8α generated a single 4kb band.

Figure 3-5 Identify resolved BAC.



3-6 Ethidium Bromide Gel showing the intact BAC. 141e18 and B1-8α were digested with EcoRI. Only the expected changes in restriction pattern were observed. A 10 kb band in 141e18 disappeared; an 8kb and a 3.5kb band appeared in the BAC B1-8α.

Figure 3-6. **Ethidium Bromide Gel showing the intact BAC**

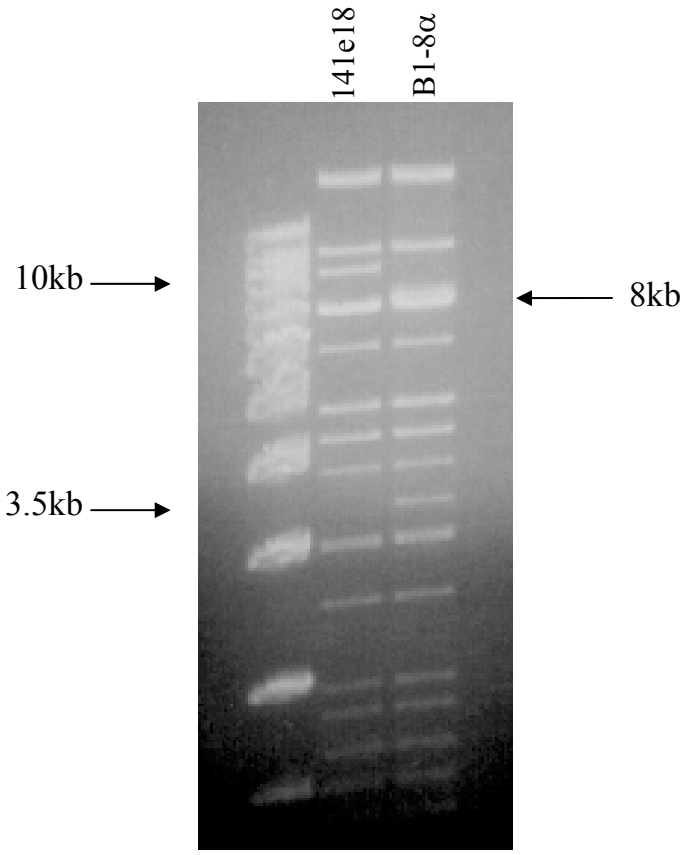
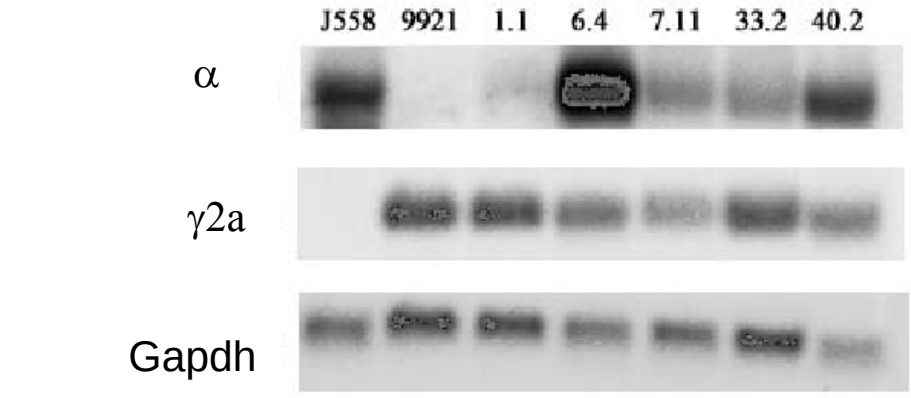


Figure 3-7. Northern blot showing the expression of B1-8α and Blxhs4. Total RNA fromB1-8α clones (A) and Blxhs4 clones (B) was isolated and used in Northern blot analysis. For the Blxhs4 clones, three independent Northern blot experiments were conducted and a representative blot is shown. In each experiment, an equal amount of RNA for each clone was loaded into two identical gels. One blot was hybridized with 300bp C_H3 region of α heavy chain, and the other was hybridized with 300bp C_H3 region of γ2a heavy chain. Subsequently, both blots were stripped and re-hybridized with normalizer GAPDH. Copy number of each clone is shown below the corresponding lanes. J558 is an IgA secreting plasmacytoma and 9921 is the IgG2a secreting parental line for all of the transformants.

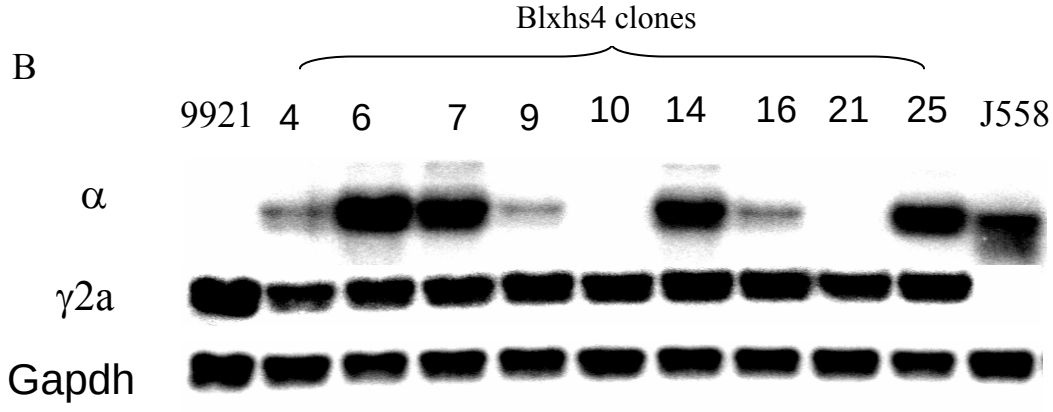
Figure 3-7 Northern blot for B1-8α and Blxhs4

A



copy number 1 4 1 1 2

B

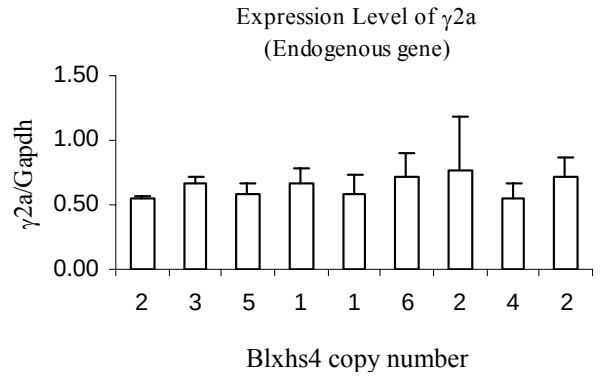


copy number 2 3 5 1 1 6 2 4 2

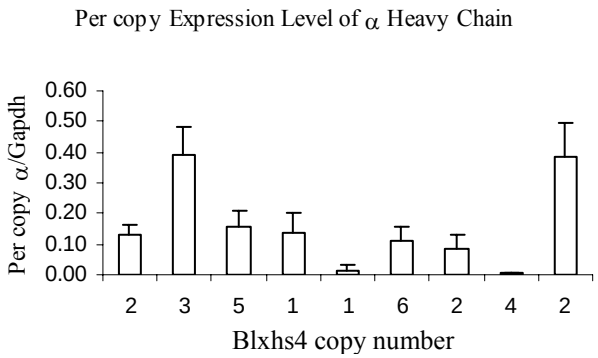
Figure 3-8. Quantitation and analysis of Blxhs4 transgene expression. (A) The expression level of γ2a heavy chain normalized to GAPDH. (B) Per copy expression level of α heavy chain normalized to GAPDH. (C) Per copy ratio of normalized α heavy chain over normalized γ2a heavy chain. (D) Linear regression showing that the expression level does not linearly correlate with copy number ($p>0.1$)

Figure 3-8. Quantitation and analysis of Blxhs4 transgene expression

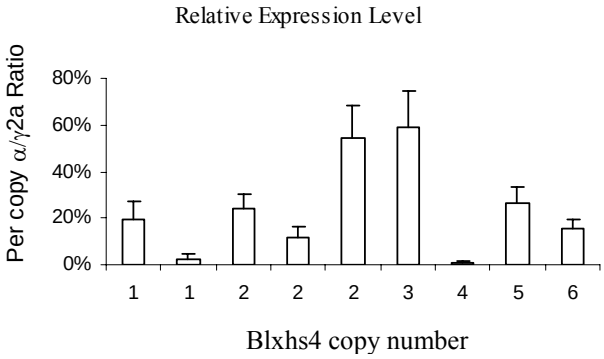
A



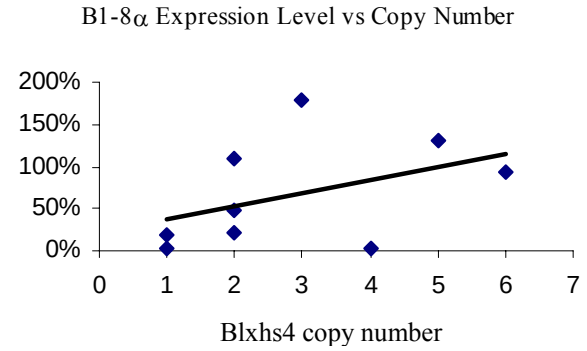
B



C



D



3.3 Synergy between the Eμ and 3' regulatory region leads to copy number-dependent transgene expression

The previous experiment showed that the 3' regulatory region present in the BAC transgene could not ensure copy-number dependent, position-independent expression and that the expression level of most of the clones was significantly lower than that of the endogenous gene. This was rather surprising because the endogenous IgH locus (γ2a of 9921) does not have Eμ, and a previous study suggested that 75%-85% of IgH transcription in Ig-secreting cells was dependent upon the 3' regulatory region even in the presence of Eμ (Gregor and Morrison 1986). There is no other element in the IgH locus known to have enhancer function. One possible explanation is that although Eμ is required to configure chromatin for high level expression in the IgH locus during development, it is not required to maintain this configuration. Because the BACs B1-8α and Blxhs4 did not go through such a developmental process, they would require Eμ to configure the chromatin for them. To test this hypothesis, we inserted Eμ between the VDJ transcription unit and Cα in BAC B1-8α, using homologous recombination as described in the previous experiment. The upstream and downstream homologous arms used were the 500bp VDJ fragment and IPAD, respectively. As shown in Figure 3-9 and Figure 3-10, both B1-8α (this time as original BAC) and the newly generated BAC which we named BEμ were digested with BamHI and hybridized with probe B consisting of VDJ sequences. B1-8α generated a 14kb band while BEμ generated a 3kb band, as was predicted by the schematic map (Figure 3-9). Ethidium bromide staining of the BamHI-digested BACs again showed that only the desired changes occurred (the 14kb band disappeared from B1-8α and a 12.5 kb band appeared in BEμ, while the addition of 3kb

was not very evident since there had already been a 3kb band in the BAC) and that no unwanted changes occurred during homologous recombination (Figure 3-11).

As was done with the BAC B1-8α, the 93kb NotI insert from BEμ was purified and isolated from the 6.7kb backbone before it was co-transfected into 9921 with neo^r and clones screened for those that had picked up BEμ. Neo^r resistant clones were screened for the presence of intact BEμ by both PCR and Southern blot. To determine copy number, genomic DNA from each clone was digested with HindIII, size-fractionated and subjected to Southern blot. As shown in Figure 3-12, the endogenous locus generated a 4.4kb band while the transgene generated a 3.7kb band. Intensities of bands were measured using phosphoimager and were used to calculate copy number. mRNA from different clones was isolated and run on a Northern blot gel. As was done with B1-8α and Blxhs4, two equivalent blots were hybridized with α and γ2a respectively, followed by GAPDH for normalization (representative data shown in Figure 3-13). The expression levels of γ2a and of α chain (per transgene copy) are shown in figures 3-14A and 3-14B. Again, the expression level of the BAC had much more variation than did the endogenous γ2a gene. The relative expression levels for different transfected clones were also calculated as the ratio of normalized α over normalized γ2a mRNA (Figure 3-14C and Figure 3-14D). As shown in Figure 3-14C, similar to Blxhs4, BEμ-transfected clones can be divided into three categories based on their per copy expression levels: High expressers with relative expression level at about 60%; intermediate expressers at about 20%, and lower expressers at about 10% of the endogenous message ($p<0.0005$).

Comparison of expression level per copy among different clones carrying BEμ and Blxhs4 (Figure 3-15) showed that there was no significant difference in the high expresser and intermediate expresser groups in both types of transfectants (both had ~20% in intermediate expressers and 60% in high expressers). The only discernable difference lies in the lower expresser groups, in which the BEμ transfected clones expressed substantially higher IgA than the "low-expresser" Blxhs4 clones. This result suggests that, although Eμ does not help increase transcription levels in permissive integration sites, nor does it help to insulate the transgene from the influence of neighboring chromatin, it helps to more effectively "open" chromatin structure in repressive chromatin environment, allowing transgenes to be expressed at these sites at substantially higher levels.

We plotted the total mRNA expression level per clone against the transgene copy number in each clone (Figure 3-14D). It is evident that overall BEμ showed a better trend in accordance with copy number than did the Blxhs4 clones (Compare Figure 3-8D and Figure 3-14D). t-test showed that the expression level of the BEμ transgene is linearly correlated with copy number ($p<0.005$).

Figure 3-9. Schematic map showing the restriction maps of B1-8α and BEμ. A 500bp VDJ fragment and IPAD (D) were used as arms for homologous recombination (hatched boxes). A 1kb fragment including Eμ was inserted between these two segments (Material and Methods). Eμ and all four 3' IgH enhancers are shown as open circles. Cε and Cα exons are shown by black boxes. Restriction enzyme: B, BamHI.

Figure 3-9. Schematic map showing the restriction maps of B1-8α and BEμ.

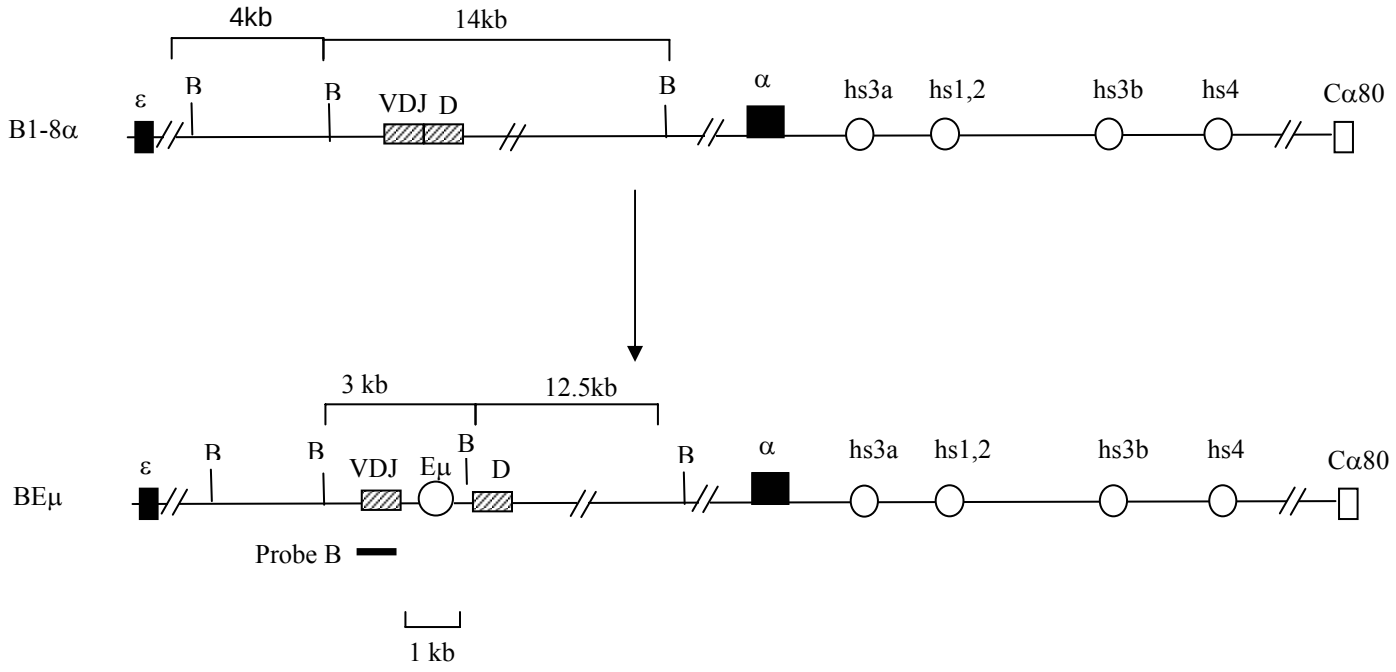


Figure 3-10 Identify resolved BEμ BAC. BACs were digested with BamHI and hybridized with VDJ (probe B in Figure 3-9). The original BAC B1-8α generated a single 14kb. The resolved BAC BEμ generated a single 3kb band.

Figure 3-10. Identify resolved BEμ BAC.

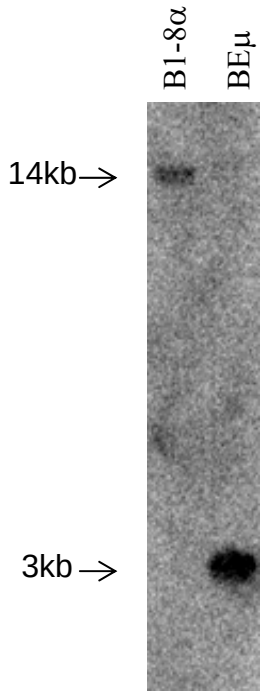


Figure 3-11. Ethidium Bromide Gel showing the intact BAC. B1-8α and BEμ were digested with BamHI. Only the expected changes in restriction pattern were observed. A 14 kb band in B1-8α disappeared; a 12.5kb and a 3kb band appeared in the BAC BEμ.

Figure 3-11. Ethidium Bromide Gel showing the intact BAC.

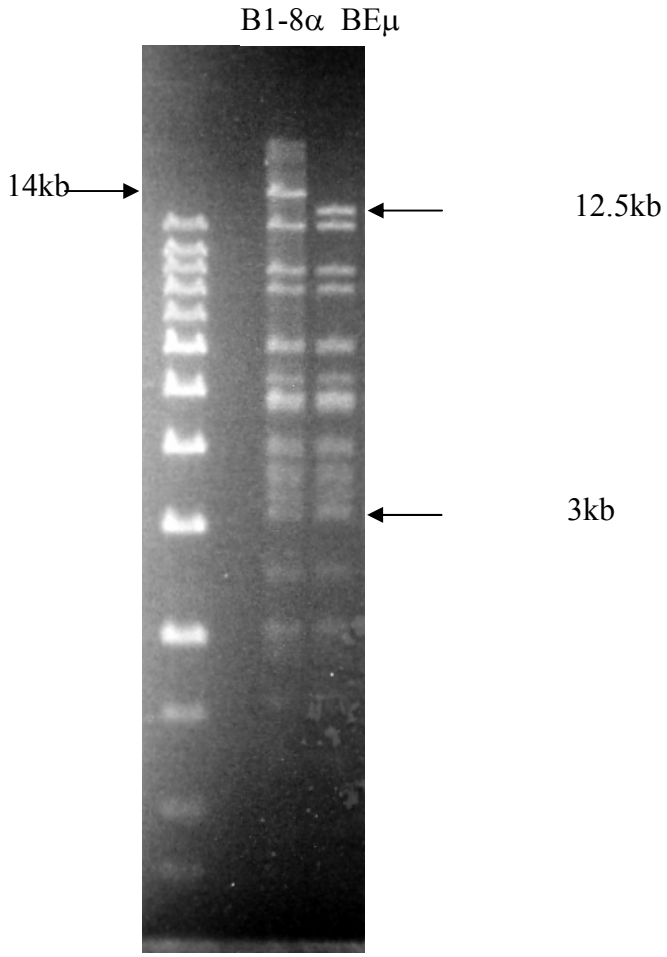
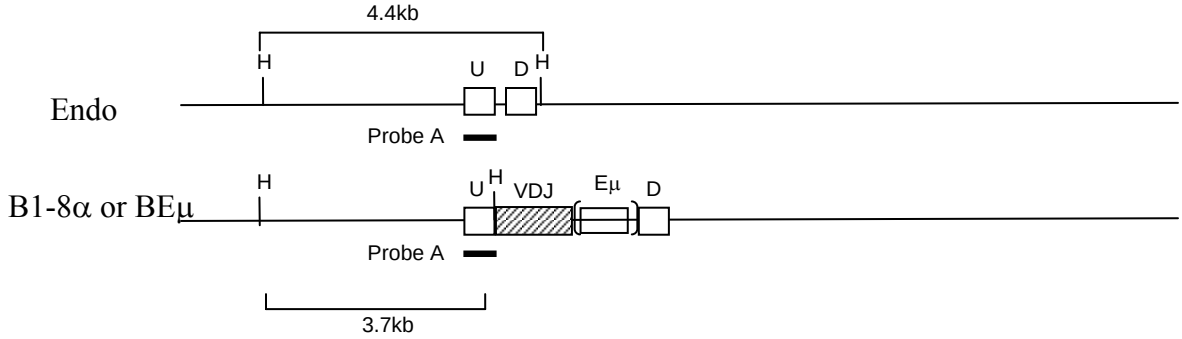


Figure 3-12. Schematic map and Southern blot used for measuring copy numbers of transgenes B1-8α, Blxhs4, and BEμ The schematic map shows the probe (probe A) used to identify a 4.4kb HindIII fragment from the endogenous locus and a 3.7 kb fragment from any of the BAC transgenes. The endogenous 4.4kb Hind III fragment is present as four copies in each cell. In this representative Southern blot, parental cell (9921) and 9 transfected cell DNAs (clone numbers 49.1 –117.1) were cut with HindIII and hybridized with probe A. Densitometry tracings of the autoradiograph were used to compare the signal for the transgene fragment vs the single-copy endogenous gene fragment. U: IPAU; D: IPAD; H: HindIII; Endo: Endogenous; Tg: Transgene.

Figure 3-12 Schematic map and Southern blot for measuring copy numbers



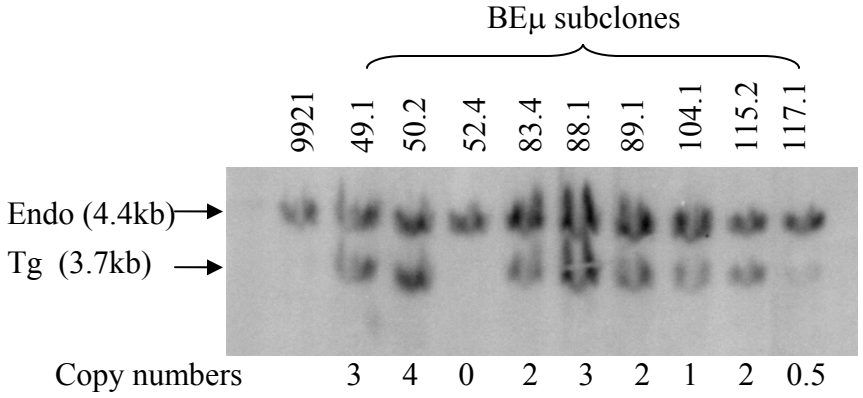
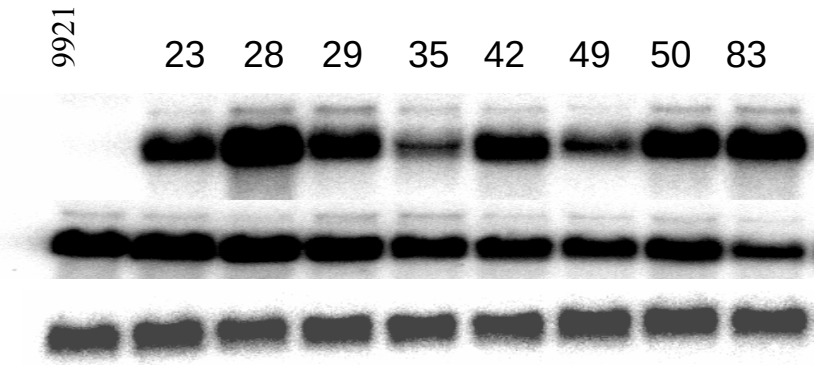


Figure 3-13. Northern blot showing the expression of BEμ transgene. Total RNA was isolated from each clone and used in Northern blot analysis. Three independent Northern blot experiments were conducted. In each experiment, equal amount of RNA for each clone was loaded into two identical gels. One blot was hybridized with 300bp C_H3 region of α heavy chain, and the other was hybridized with 300bp C_H3 region of γ2a heavy chain. Subsequently, both blots were stripped and re-hybridized with normalizer GAPDH. Copy number of each clone is shown below corresponding lane.

Figure 3-13. Northern blot showing the expression of BEμ transgene



copy number	1	7	7	3	5	3	4	2

Figure 3-14. Quantification and analysis of BEμ transgene expression

. (A) The expression level of γ2a heavy chain normalized to GAPDH. (B) Per copy expression level of α heavy chain normalized to GAPDH. (C) Per copy ratio of normalized α heavy chain over normalized γ2a heavy chain. (D) Linear regression showing that the expression level linearly correlates with copy number ($p<0.005$)

Figure 3-14. Quantification and analysis of BEμ transgene expression

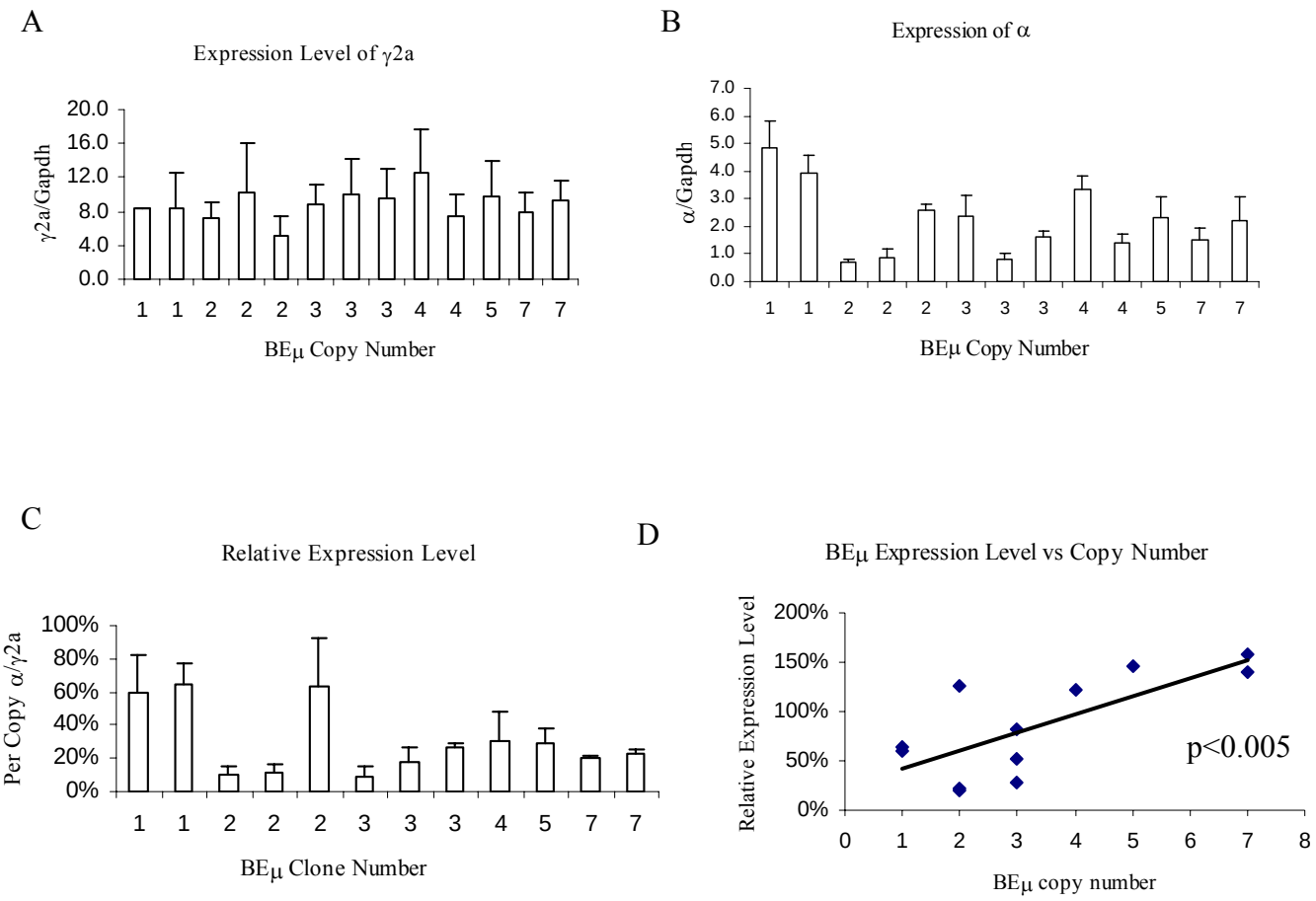
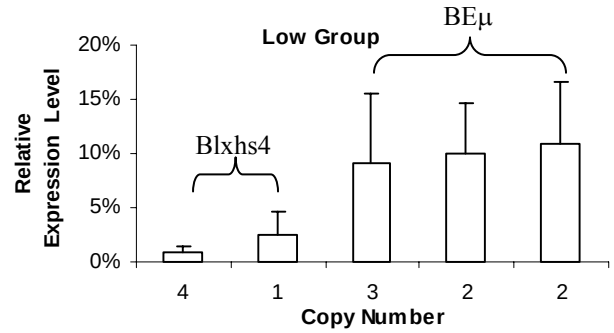
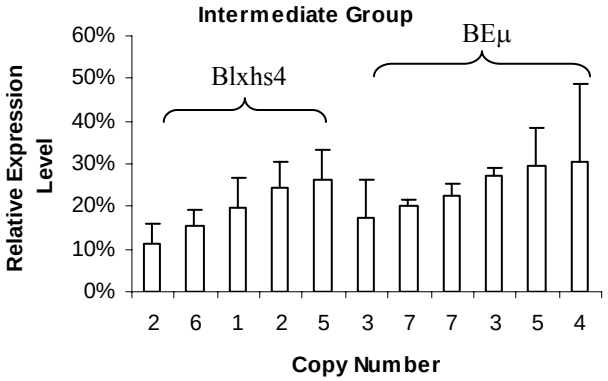


Figure 3-15 Compare expression levels between Blxhs4 and BEμ transgenes in groups. Per-copy expression levels of Blxhs4 and BEμ transgenes are compared side-by- side, grouped into low, intermediate and high- levels, which indicate non-permissive, semi-permissive, and permissive integration sites, respectively.

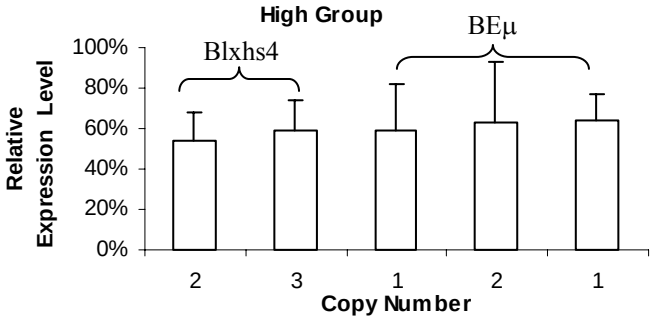
Figure 3-15. Compare expression levels between Blxhs4 and BEμ transgenes in groups



non-permissive Chromatin



Semi-permissive Chromatin



permissive Chromatin

3.4 hs4 deletion alone does not eliminate transgene expression

In previous experiments, we showed that the hs3b/hs4 enhancer pair was critical to expression of an IgH mini-locus in Ig-secreting cells (Shi and Eckhardt, 2001). In those experiments, as well as in the present ones, expression of the locus was established before the enhancer pair was deleted. Interestingly, a study involving deletion of the same enhancer pair within the IgH loci of genetically modified mice established a critical role for these enhancers in class-switch recombination (Pinaud et, al 2001). One of the questions that arose from both studies was whether the effects seen were due to deletion of hs3b alone, to deletion of hs4 alone, or were achievable only after deletion of both of these regulatory sequences. Several findings implicated hs4. First, hs4 has several attributes that distinguish it from the other enhancers. In transient transfection assays, hs4 is the most robust of the 3'IgH enhancers in mature B cell lines and in plasmacytoma lines (Ong et al., 1998). It is also the only one of the 3' IgH enhancers that shows DNaseI hypersensitivity in both pro-B and pre-B cells; the other enhancers become DNaseI hypersensitive later, at the surface Ig+ cell stage (Giannini et al 1993; Madisen and Groudine 1994; Garrett et al. 2005). Secondly, there was reason to doubt that deletion of hs3b alone could lead both to loss of IgH gene transcription in the IgH mini-locus of the cell line studies and to a dramatic effect on heavy chain class switching in mice. hs3b is almost identical to hs3a, and deletion of the latter had no effect on heavy chain class switching as measured in genetically modified mice (Manis et al., 1998) nor on transcription of an IgH mini-locus in Ig-secreting cells (Shi and Eckhardt , 2001). This might reflect a functional redundancy between these two elements, further

implicating hs4 as the critical element in the hs3b/hs4 deletions that gave rise to the observed phenotypes.

To test this hypothesis, we modified the B1-8α BAC to allow for deletion of hs4. Because the per-copy expression level differed among independent B1-8α transformants, we elected to generate several independent transformants and test the effect of hs4 deletion after stable integration of the BAC into the genome of the recipient cells. In this way, each transformant provides information about BAC gene activity at a single site of integration, both before and after hs4 deletion.

As shown in Figure 3-16 and Figure 3-17A, hs4 was deleted from B1-8α by homologous recombination in bacteria. The homologous arms used in this modification were HS4U and HS4D, respectively. Because the 2.5kb DNA sequence between HS4U and hs4 contains highly repetitive sequences, we and others were not able to sequence and clone this fragment, let alone use it as an homologous arm. As a result of this homologous recombination, this 2.5kb DNA was also deleted resulting in a total of 3.9kb being removed from this region of the BAC. DNA was cut with BamHI and hybridized with either HS4D (Figure 3-17A) or HS4U (Figure 3-17B). The Southern blot in Figure 3-17A clearly showed that hs4 was deleted from the BAC B1-8α as evident by the shifting of the 2kb to a 1.8kb BamHI fragment (HS4D probe, refer to maps in Figure 3-16). BAC Blxhs4 was generated by inserting floxed hs4 followed by a cDNA tag consisting of β-globin cDNA sequence in place of the original hs4. This DNA tag would enable us to detect differences, if any, in DNaseI hypersensitivity at hs3b before and after hs4 deletion (Figure 3-16). Southern blot analysis confirmed successful construction of Blxhs4 (Figure 3-17B). We then transfected 9921 with Blxhs4, screened for clones that

contained the entire BAC insert and measured the IgA vs IgG2a mRNA level using Northern blot. As described earlier, we found that the transfected cell lines expressed IgA at levels similar to those of B1-8α (Figure 3-7A), demonstrating that the BAC Blxhs4 is equivalent to B1-8α in terms of transcription and that the 2.5kb of DNA deleted in this BAC does not play a role in transcription. hs4 was then deleted from the BAC transgene by transiently transfecting these cell lines with a plasmid that contains genes that encode both GFP and CRE, respectively. Green cells, which supposedly had taken up the plasmid, were sorted into 96well plates containing culture media at one cell per well frequency by fluorescence activated cell sorter (FACS, Vantage^{TM}, Becton-Dickinson and screened by PCR for clones that had deleted hs4 (Material and Methods; Data not shown).

Clones that harbored deletion of hs4 as shown by PCR were further demonstrated to have undergone this deletion by Southern Blot (Figures 3-18 and Figure 3-19). Genomic DNA was first digested with HindIII and then hybridized with hs3b. The endogenous locus generated a fragment of 8kb. The Blxhs4 transfectants with intact hs4 generated a 5.5kb fragment while the hs4-deletion clones generated a 4kb band, as predicted (Figure 3-18 and Figure 3-19). By this Southern blot it was clearly shown that hs4 had been deleted from all five independent Blxhs4 transfectants. DNA surrounding the loxP site was amplified by PCR and sequenced to show that hs4 was indeed deleted through CRE-loxP mediated recombination and that there was no aberrant change in the vicinity of the loxP site (Data not shown).

We sought to see whether there was an effect on transcription from the transgene upon deletion of hs4. Total RNA was isolated from clones with or without hs4

respectively and hybridized with a Cα probe and the results normalized with the endogenous γ2a chain message (Figure 3-20). To our surprise, hs4 deletion from the BAC had no dramatic effect on the transcription levels of the transgene. The relative expression of α heavy chain mRNA from the transgene over γ2a heavy chain mRNA reduced in a range from 20% to 50%, while the relative expression of one clone (6.1) increased by 40% (Figure 3-21).

Therefore hs4 is not essential for maintaining high level expression of the transgene in an already established open chromatin environment. All hs4 can do at best is to contribute to 20%~50% transcription enhancing activity to the 3'IgH regulatory region. This result, combined with our previous experiments, suggests that either Hs3b alone is the critical element for high level transcription or the loss of hs4 can be compensated by some other element (presumably hs3b) in the 3'IgH regulatory region. Experiments are underway to test this hypothesis.

Figure 3-16. Restriction maps of BACs B1-8α and B1-8αΔHS4 and Blxhs4. HS4U and HS4D (hatched boxes) are used as homologous arms to delete hs4 from B1-8α and insert floxed hs4 (hs4 flanked by loxP sites). hs3a and hs1,2 are shown as small circles. The deletion giving rise to B1-8αΔHS4 was 3.9kb, which included a 1.4kb region containing hs4 and 2.5kb upstream (25810 to 27324 of sequence AF450245. Note, in AF450245, some sequence is not included, because of sequencing difficulty). The insertion giving rise to Blxhs4 was 1.4kb hs4 fragment with 300bp of DNA flanking hs4 with embedded loxP sites, and a 420bp Tag sequence. Blxhs4 differs from B1-8α both by the insertion of loxP sites, insertion of Tag DNA consisting of a portion of β-globin cDNA sequence, and deletion of a 2.5kb DNA sequence upstream of hs4. hs3b and hs4 are shown as open boxes. black box ,Cα exon; Black triangles, loxP sites; Tag, a DNA tag consisting of portion of β-globin cDNA sequence. Restriction enzyme: B, BamHI.

Figure 3-16. Restriction maps of BACs B1-8α, B1-8αΔHS4 and Blxhs4

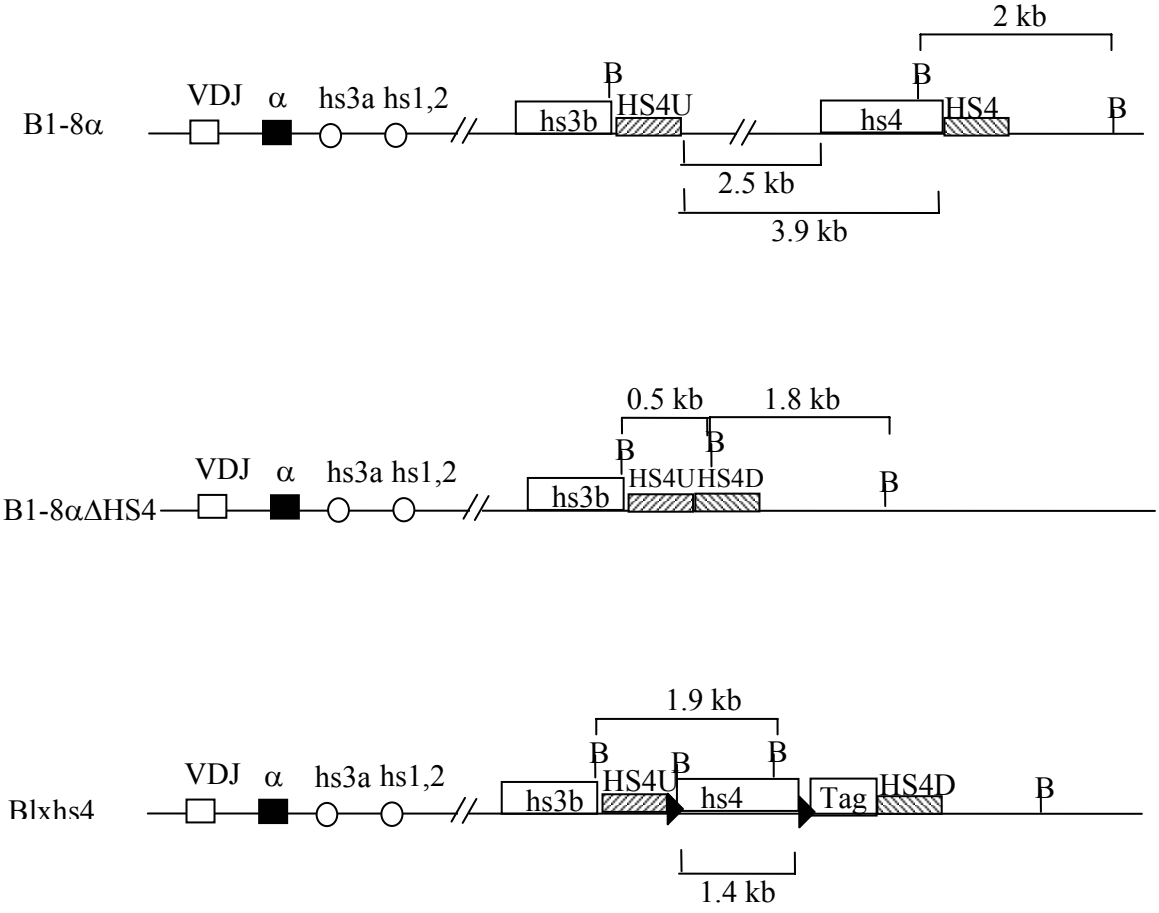


Figure 3-17 Southern blots showing the construction of BACs B1-8αΔHS4 and Blxhs4. (A) BACs B1-8α and B1-8αΔHS4 were digested with BamHI and hybridized with HS4D. B1-8α generated a 2kb band while B1-8αΔHS4 generated a 1.8kb band. (B) BACs B1-8αΔHS4 and Blxhs4 were digested with BamHI and hybridized with probe HS4U. B1-8αΔHS4 generated a 0.5kb band, while Blxhs4 generated a 1.9kb band.

Figure 3-17. Southern blots showing the construction of BACs B1-8αΔHS4 and Blxhs4

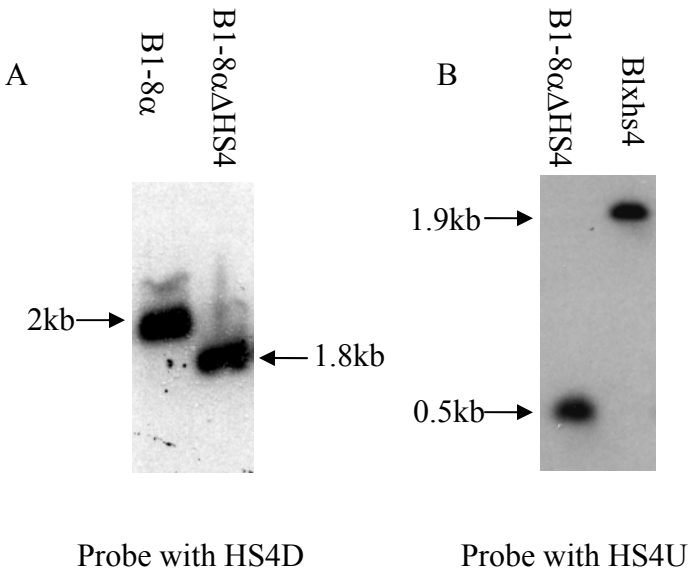


Figure 3-18 Restriction maps of the 3'IgH region of the endogenous locus and the BAC transgene before--Blxhs4 and after deletion of hs4—Blxhs4Δ.

The lengths of HindIII fragments detected by hs3b are shown in the brackets above the map. hs3a and hs1,2 are shown as small circles. Black boxes Cγ2a and Cα exon; Open boxes, rearranged VDJ fragment; Black triangles, loxP sites; Tag, a DNA tag consisting of a portion of β-globin cDNA sequence. Restriction enzyme: H, HindIII.

Figure 3-18 Restriction maps of the 3'IgH region of the endogenous locus and the BAC transgene before--Blxhs4 and after deletion of hs4--Blxhs4Δ.

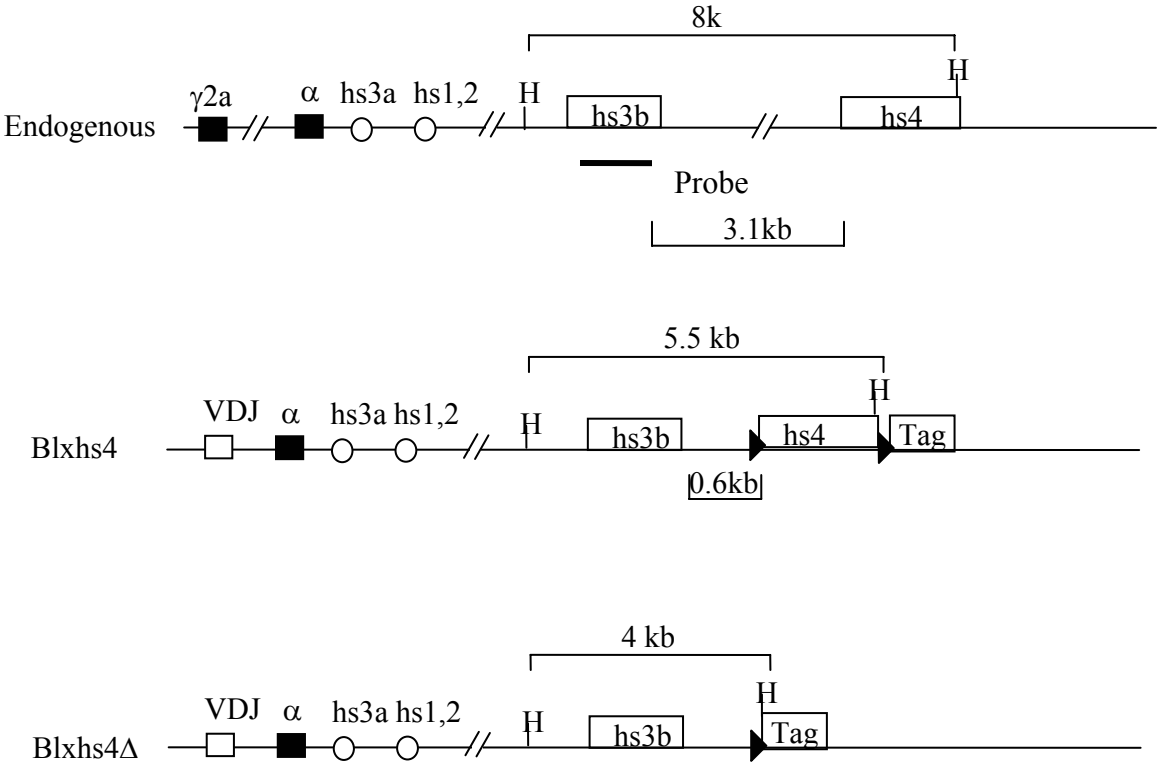


Figure 3-19. Southern blot showing the screening for clones with hs4 deletion. Clones with and without hs4 in the transgene are shown side by side. Minus sign, without hs4. Plus sign, with hs4 present. Genomic DNAs were digested with HindIII and hybridized with hs3b. This probe also hybridizes to a 1.5 kb HindIII band, which contains hs3a, but after gel transfer to nylon, this region of the nylon was removed before hybridization. I

Figure 3-19. Southern blot showing the screening for clones with hs4 deletion.

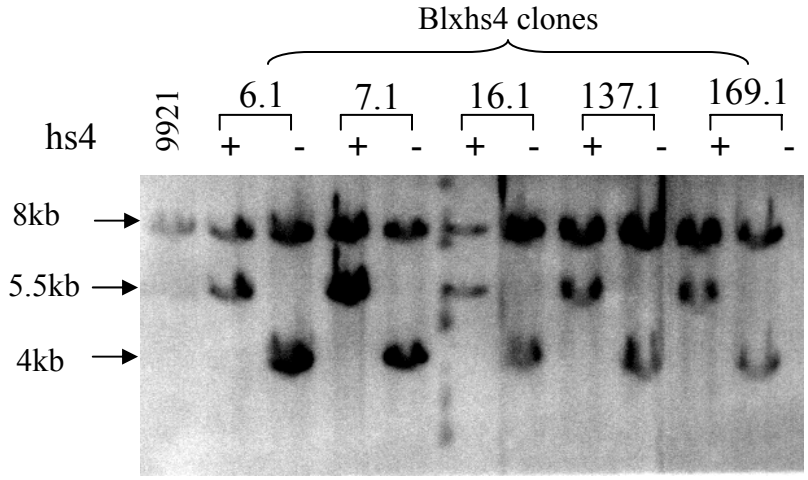


Figure 3-20. Northern blot showing the expression of transgenes with and without hs4. Total RNA was isolated from clones that retained or had deleted hs4 and loaded side by side and subjected to Northern blot analysis. Three independent Northern blot experiments were conducted by hybridizing with 300bp C_H3 region of α heavy chain followed by hybridizing with 300 bp γ2a heavy chain and GAPDH for normalization. 9921 is the parental line used for these transfection experiments (expresses γ2a). J558 is an IgA-secreting plasmacytoma.

Figure 3-20. Northern blot showing the expression of transgenes with and without hs4

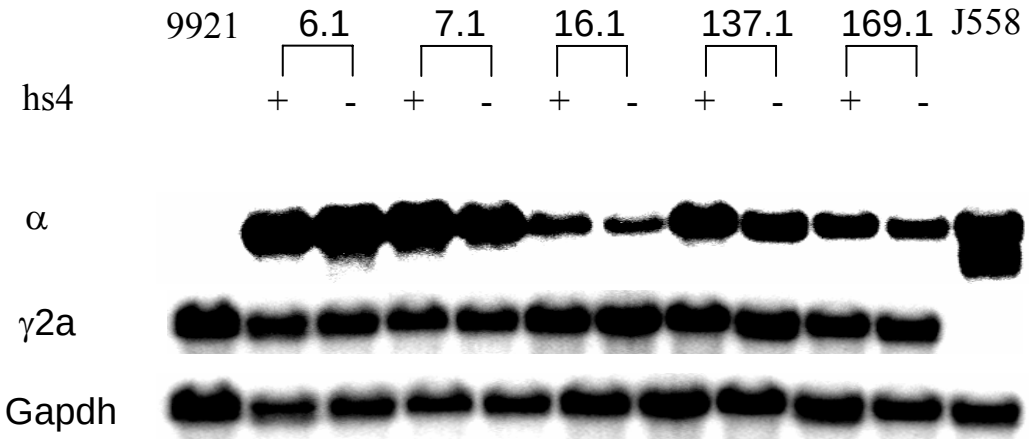
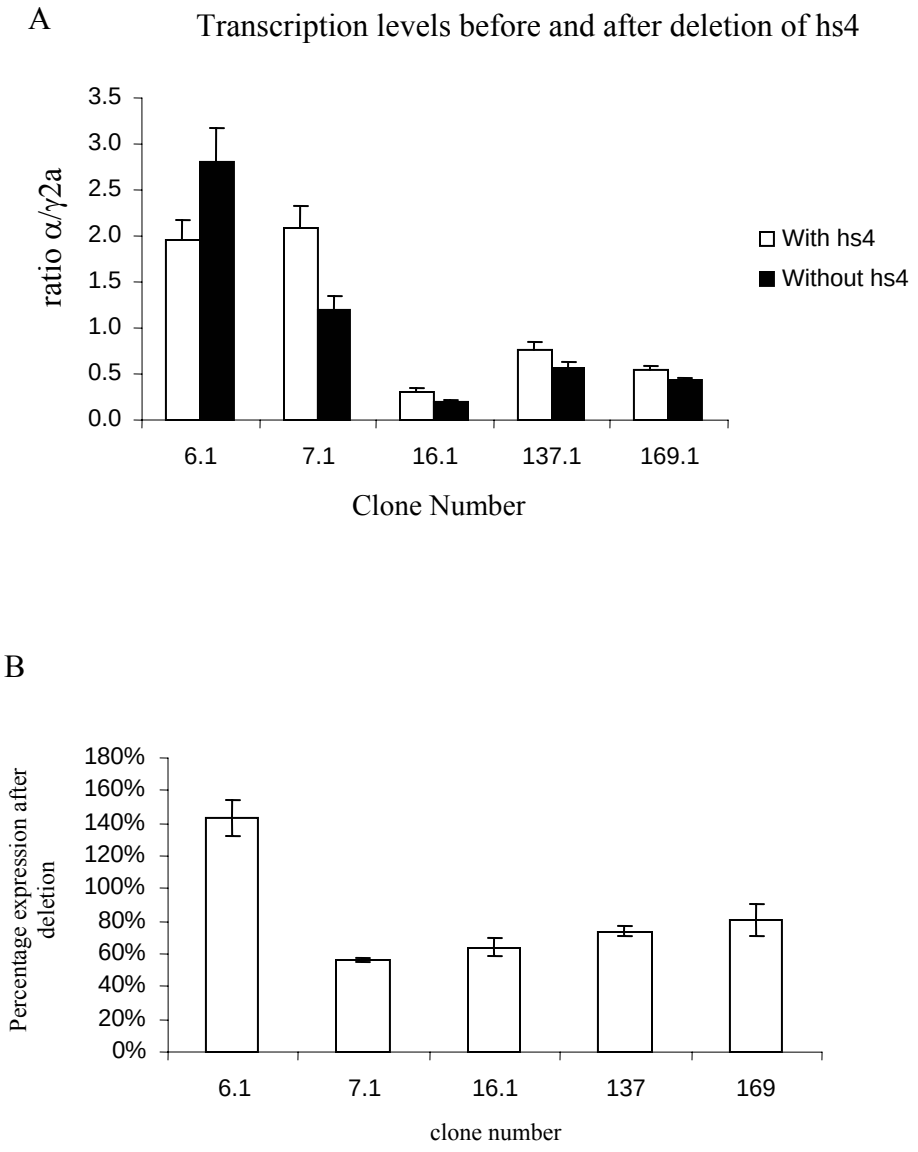


Figure 3-21. Expression level of the BAC transgene before and after deletion of hs4. Three independent northern blots were conducted. mRNA levels of α transgene and endogenous γ2a gene were quantified with phosphorimager. (A) Relative expression level of the transgene over the endogenous gene was calculated for clones with and without deletion of hs4. (B) The percentage α expression of clones after deletion of hs4.

Figure 3-21. Expression level of the BAC transgene before and after deletion of hs4



Chapter 4 General Discussion

4.1 The rationale behind studying LCR activity and various standards for defining LCRs.

Locus Control Region (LCR) is operationally defined in transgenic mice as the combination of cis-elements that confer upon transgenes both copy number-dependent, and position-independent expression in a tissue and stage specific manner. Our impetus to identify an LCR within a locus roots from the need to know all the cis-elements that are necessary and sufficient to confer the physiological (natural) level of transcription at the endogenous locus. Once such an LCR is defined, it would provide a framework for understanding the mechanisms of transcription of tissue-specific genes and manipulating gene expression in different developmental stages of cells, which has both theoretical and medical implications (See significance section in Chapter 1 Introduction). As more data come along, and as synergy and redundancies among cis-elements are discovered, the operationally defined LCR seems not able to meet all the criteria we want it to meet. Therefore, the concept and the usage of LCR are vague in the literature and are evolving. Nevertheless, since any new idea or model about LCR awaits data to gratify, the operationally defined concept is used in this thesis.

The concept of LCR was first derived from early studies of the β-globin locus. It was found that the expression of an LCR linked gene was in strict linear relationship with copy number of the transgene (Grosveld et al., 1987), indicating that the expression of the transgene is not influenced by neighboring chromatin or regulatory sequences, and thus is position independent. Later, LCRs of many other loci have been discovered. For a

list of these LCRs, please refer to a review on LCRs (Li et al., 2002b). It should be noted that, among these newly defined LCRs, many of them confer transgene copy number dependent expression only in a statistical sense. That is, although the average expression level in any individual mouse or cell line is correlated with transgene copy number with high statistical confidence, the expression level of clones or mice with the same copy number could differ as much as 10 fold. These somewhat conflicting observations suggest that, while expression of the transgene is largely not influenced by neighboring chromatin, the transgene is not completely insulated from the influence of neighboring regulatory elements. As discussed in the Introduction section, the prototypical LCR is that in the β-globin locus, which confers strictly copy-number dependent expression. LCRs that confer transgene expression but not necessarily in a strictly copy number-dependent manner are defined as **partial LCRs**.

Such categorization brings up the question, how physiologically relevant are this prototypical LCR and partial LCRs—that is, how much fluctuation of expression of certain genes is there in reality? Another closely related question is, how much fluctuation in the expression of certain genes can be tolerated by an organism? It can be postulated that a strict LCR is required for normal gene expression in some loci, because subtle changes in the expression level of some genes could have devastating physiological consequences. For example, as little as a 15% increase in the expression level of CD19 could cause autoimmune disease in mice (Sato et al., 2000). This result was correlated with the fact that patients with multiple sclerosis have 20% higher expression of CD19 than the normal level, suggesting that an LCR controlling CD19 must be able to regulate expression of the CD19 gene with sharp precision. On the other

hand, expression of some other genes may not require such precision. For example, the requirement for accuracy of expression level of the Ig heavy chain may not be so stringent. Small fluctuations in antibody production are unlikely to cause a dramatic effect. In addition, from the point of view of antibody production, the number of antibodies produced depends not only on the transcription level of the heavy chain in individual plasma cells, but also depends on, at least, (1) post-transcriptional controls, e.g mRNA stability regulation; (2) the efficiency of secretion of individual plasma cells, (3) the types and strength of antigen and T cell stimulation, and (4) the number of plasma cells proliferating. Thus multiple mechanisms are involved in regulating the serum level against certain antigen to guarantee that enough antibodies are produced. It is possible that a large window exists for fluctuation of the transcription of the Ig heavy chain.

4.2 9921 as a suitable model for studying the functions of regulatory regions in the IgH locus.

In this study, a plasmacytoma cell line 9921 was used as a representative of Ig secreting cells, or plasma cells. Plasma cells are terminally differentiated and no longer proliferate. These characteristics determine that plasma cells directly isolated from mice are not suitable for transfection experiments. In addition, it is technically difficult to isolate plasma cells directly from mice. Therefore we did not use direct isolation of plasma cells to conduct this study.

Nevertheless, cell line 9921 can represent most of the behaviors of plasma cells related to the transcriptional regulation of the IgH genes. It secretes immunoglobulin composed of a γ2a heavy chain and a κ light chain. It also expresses important transcription factors present in plasma cells. They are BSAP, Oct2, Oct1, OcaB and PU.1.

What might yield 9921 unsuitable for this study are the translocated *c-myc* genes deregulated by the 3'IgH regulatory region. *c-myc* is a proto-oncogene found to be over-expressed in various cancers, including human Burkitts lymphomas, in which *c-myc* has undergone similar translocations as seen in the cell line 9921. *c-myc* has numerous putative downstream genes, most of which are involved in cell cycle, apoptosis, and metabolism (Dang., 1999). Although none of them are directly involved in transcription regulation of the immunoglobulin genes, it cannot be ruled out that the ability of *c-myc* to interact with TATA box binding protein (TBP) might affect the outcome of transfection assays (Dang., 1999). However, neo^r replacement of hs1,2 diminished the expression of the endogenous γ2a heavy chain in 9921(Lieberson et al., 1995). The hs3b/4 double deletion also abolished transcription of a γ2b transgene (Shi and Eckhardt., 2001). These results suggest that in 9921, the transcription of immunoglobulin genes is still dependent on the 3'IgH regulatory region. Therefore, influence of the *c-myc* proto-oncogene on the transcription of the BAC transgene, if any exists, would be secondary or minimal, and would have to be mediated through the cis-regulatory elements of the IgH locus.

4.3 The intact 3'IgH RR constitutes a partial LCR, and its combination with Eμ also constitutes a partial LCR.

4.3.1 The dispute over whether the 3'IgH RR constitutes an LCR.

Molecular immunologists have been studying for a decade to answer the question whether the 3'IgH regulatory region behaves as an LCR, that is, whether the 3'IgH regulatory region could confer transgene copy number-dependent, position independent expression. The hypothesis is, although every one of the 3'IgH enhancers is relatively weak, synergies among them enable the whole region to overcome position

effect completely. All of these studies largely agreed that the combination of 3'IgH enhancers enabled the expression of a transgene under the control of any promoter in almost every integration site (Madisen and Groudine 1994; Chauveau et al., 1999; Shi et al, 2001). These results suggest that the 3'IgH regulatory region can counteract the spreading of heterochromatin and open the chromatin where the transgene resides in most cases. Part of this ability has been attributed to the ability of 3'IgH enhancers to modify histones (Madisen et al., 1998). However, only one of the above studies, which utilized three of the four 3'IgH enhancers, was able to show that the transgenes under the control of 3'IgH enhancers were expressed in a copy number- dependent manner with statistical significance (Madisen and Groudine 1994). The other two studies with transgenic mice (Chauveau et al., 1999) and with cell lines (Shi et al., 2001) showed that this was not the case. Therefore, it has been widely believed that the combination of the four 3'IgH enhancers, hs3a, hs1,2, hs3b, and hs4, constitutes a partial LCR. However, it is questionable whether these studies with the four 3'IgH enhancers juxtaposed together could completely mimic the physiological activity of the 3'IgH regulatory region. On one hand, this juxtaposition may undermine the possible LCR activity possessed by the 3'IgH regulatory region by omitting unknown control elements in this region and/or by disrupting possible secondary structure formed by the palindrome centered about hs1,2. On the other hand, this juxtaposition may artificially display LCR properties that do not belong to the 3'IgH regulatory region by placing them in close vicinity that may enhance the synergy among these enhancers. It should also be noted here that the traditional characterization of enhancers as orientation and distance-independent elements was based

on transient transfection assays which neither separate the enhancers from promoters more than a few kilobases away nor reflect physiological chromatin structure.

4.3.2 The 3'IgH RR constitutes a partial LCR; Comparison to earlier studies.

To study the possible LCR activity of the 3'IgH regulatory region in a more natural context and yet avoid the difficulties rooted in knock out enhancers in the endogenous IgH locus, we developed a strategy using BAC as the carrier for transgenes. The reporter BAC B1-8α contains all the known and unknown 3'regulatory elements in their natural positions and orientations. It was later found out that the IgH locus likely ends 60kb downstream of Cα, showing that this and other BACs we constructed contain the entire 3'RR (Garrett et al., 2005). The expression level of the transgene in this BAC was not copy number dependent (a linear regression gives a p value that is greater than 0.1), suggesting that the entire 3'IgH regulatory region is a partial LCR. This result looks similar to that observed previously in mini-constructs that contained a cluster of 3'IgH enhancers (Shi and Eckhardt, 2001; Chauveau et al., 1999), but is different from that observed in the first study on the LCR activity of the 3'IgH region (Madisen and Groudine, 1994).

This discrepancy could be due to different promoters used in these studies. The Madisen and Groudine study used the *c-myc* promoters while the other two and this study used IgH promoters. Although no study has been conducted to directly compare the strength of the two kinds of promoters, comparison between data in two different labs shows that the 3'IgH enhancers could enhance transcription from a *c-myc* promoter at least one order of magnitude higher than that from a V_H promoter (Madisen and Groudine 1994; Ong et al., 1998). Because both promoters and enhancers are clusters of transacting

factor binding sites, it is possible that the *c-myc* promoter acts like another component of an artificial LCR and that the strong synergy between the *c-myc* promoter and the 3'IgH enhancers could not only overcome partially but also override completely position effect variegation (PEV). Studies using a *c-myc* transgene under the control of the 3'IgH enhancers from collaboration between our lab and that of Dr. Siegfreid Janz lab at NIH could potentially help resolve this issue.

4.3.3 Combination of Eμ with 3'IgH RR constitutes a partial LCR; Its comparison to 3'IgH RR alone.

The pattern of transgene expression among BEμ transfected clones was strikingly similar to that observed by Madisen and Groudine, and the expression level of the lower expressing group is significantly higher than that of the Blxhs4 transformants. Although the transgenes are entirely different (IgH gene vs c-*myc*), the ratio of the expression levels among the three groups is exactly the same (6:2:1) (with ANOVA test $p<0.0005$ and Brown-Forsythe test for robust $p<0.0005$).

Linear regression showed that the BAC BEμ is expressed in a copy number dependent manner statistically ($p<0.005$, $r=0.73$), while Blxhs4 is not ($p>0.1$ $r=0.43$), suggesting that Eμ, in combination with the 3'IgH RR, constitutes an LCR.

It is unlikely that Eμ alone was responsible for the statistically significant correlation between copy number and expression, since many early studies with transgenes with Eμ alone showed that these transgenes were not expressed in a copy number dependent manner (Moroy et al., 1990; Dildrop et al., 1989; Grosschedl and Marx 1988; Guglielmi et al., 2003a).

The results obtained with BAC BEμ are not in agreement with those seen with a transgene in which Eμ, in combination with the cluster of all four 3'IgH enhancers, was not able to confer copy number-dependent expression of a β-globin gene with a V_H promoter in a B cell line (Guglielmi et al., 2003b,c). This discrepancy may arise from different cell lines used. In the earlier study, pre-B and mature B cell lines were used, while a plasmacytoma cell line was used in this study. Given this result, however, we are obliged to offer an explanation for the fact that many plasmacytoma cell lines have lost Eμ but retained their expression level (Anguilera et al, 1985; Eckhardt and Birshtein 1985; Zaller and Eckhardt 1985). One possibility is that Eμ is required for the establishment of LCR activity but not to maintain it. In order to test this hypothesis, a floxed Eμ (loxP flanked Eμ) could be inserted into BAC B1-8α in place of Eμ. Following the establishment of transfected clones, Eμ could be deleted through CRE-loxP mediated recombination. By comparing the expression levels before and after the deletion of Eμ, we should be able to distinguish the above two possibilities.

What in Eμ is different from B1-8α and Blxhs4 that confers this copy number dependence? As has been described in the chapter 1—introduction, the 1kb fragment of Eμ consists of a 220bp core enhancer and two matrix attachment regions that flank it. The core contains many transcription factor binding sites but could only increase local accessibility to the transgenes, while the addition of MARs extended this accessibility over a longer range (Jenuwein, et al., 1993; Jenuwein et al., 1997). It has recently been shown that MARs affected acetylation of histones H3 and H4 at Eμ enhancer distal nucleosomes. (Fernadez et al., 2001). These studies underscore the importance of MARs in directing long range chromatin accessibility and, possibly, long range interactions

between promoters and enhancers. In order to test whether the MARs in Eμ mark the differences between the transcription profile of Blxhs4 (and B1-8α) and BEμ, a BAC with MARs inserted between VDJ and Cα has been constructed to test whether MARs can exert the function of Eμ.

As another way to evaluate the contribution of Eμ and of the IgH cis-elements in the absence of Eμ beyond VDJ recombination in B cell development, a rearranged VDJ fragment B1-8 has been inserted into the endogenous locus and at the same time replaced Eμ. This experiment will not only answer the question whether Eμ is required for recombination events, transcription and somatic hypermutation in later stages of B cell development, but also will allow us to elucidate the contributions of other elements and their redundancies with Eμ in systematic follow-up experiments.

4.4 The transgene expression in BEμ is not strictly copy number dependent.

4.4.1 Eμ and the 3'IgH RR together, do not constitute a full LCR.

As mentioned above, the expression levels of BEμ transfected clones are linearly correlated with copy number with statistical significance. The correlation coefficient is merely satisfying (r=0.73). The expression levels of clones with the same copy number vary as much as 6 fold. In addition, that BEμ clones can be divided into three categories based on their per copy expression levels, suggests that Eμ plus the 3' regulatory region still could not completely override the position effect in independent integration sites. Are there control elements in the IgH locus that we are missing, the inclusion of which into the BAC transgene could yield strict copy number dependence?

One possibility is that BAC BEμ lacks a 5'insulator. Insulators are cis-elements that block the influence of neighboring regulatory or chromatin structure. According to the pioneering studies by Felsenseld's group on the chicken β-globin locus, insulator function seems to have two aspects: enhancer blocking activity and chromatin effect barrier. These two activities are separable. The chicken β-globin insulator (5'HS4), was dissected into five distinct fragments (FI~FV). FII contains a binding site for CTCF and was shown to be both necessary and sufficient to block interactions between promoter and enhancers (Recillas-Targa et al., 1999). However, this fragment is neither necessary nor sufficient for the chromatin barrier function of the β-globin insulator. The chromatin barrier function in the chicken β-globin locus is carried out by the combination of the fragments FI, FIII, FIV and FV of the insulator, measured by the ability to confer transgene copy number dependent expression. (Recillas-Targa et al., 2002).

It has recently been shown that the 3'IgH regulatory region contains a putative insulator 6kb downstream of hs4.This insulator possesses CTCF binding sites and has been characterized as an enhancer blocker in a transient transfection assay. Whether it can act like a chromatin effect barrier awaits further experimental data (Garret et al., 2005). To date, no cis-element with insulator function has been found in the vicinity of the 5' end of the IgH locus. However, the replication timing of various V_H region genes are different and is developmentally controlled (Zhou et al., 2002). Recently, a cluster of three DNaseI hypersensitive sites has been located 30kb upstream of the most 5' end of the most 5'IgH V_H gene, and were shown to bind PU.1 and E2A in a B cell specific manner (Pawlitzky et al., 2004). These studies suggest that unknown cis-elements exist at the 5' end of the IgH Locus. The 5' end of the IgH locus is 4 megabases from the

telomere. At least three non-IgH genes exist between the 5'end of the IgH locus and the telomere (NCBI Mapviewer). The complexity of B cell specific DNA rearrangement and transcription suggests that the IgH locus is structurally and functionally separated from these non-IgH genes, probably through insulators, which are able to block the influence of the IgH locus on the expression of these non-IgH genes and prevent the spreading of the suppressive effect of the telomere to the IgH locus.

Alternatively, the lack of strict copy number-dependent expression could be accounted for by the lack of cis-elements in the endogenous locus other than insulators. Several matrix attachment regions (MARs) have been found in the intervening sequences between δ and γ3, γ3 and γ1, and γ1 and γ2b constant regions (Cockerill 1990). A large portion of these attachment regions binds to the matrix more robustly than Eμ does. Although no functional studies have been conducted with these MARs, it is possible that, during development, these regions help to relocate the IgH locus to subnuclear compartments which allow opening of the locus to establish high level gene expression but not to maintain it. Consequently, following such activation, as happened in the cell line 9921, MARs could be lost without causing a dramatic effect on transcription.

Recently, a 230 kb BAC transgene encompassing J_H through the most 3' IgH region was injected into transgenic mice. The serum level IgG1 after immunization measured by ELISA was shown to be strictly copy number dependent, suggesting that this BAC contains all the required cis-elements for a typical LCR. (Dunnick et al., 2004). However, the γ1 constant region seems to be different from any other constant region in that neo^r replacement or insertion in various positions in the IgH locus dramatically reduced CSR to isotypes other than γ1, suggesting that CSR to and transcription of the γ1

gene is somewhat independently regulated (Cogne et al., 1994; Manis et al., 1998). Therefore, the γ1 gene might be regulated independently of Eμ and the 3'IgH RR and the conclusion obtained from this result may not be applicable to transgenes.

To address this possibility that the BAC BEμ lacks one of the components possessed by the endogenous locus (either an insulator or MARs), the same rearranged VDJ could be inserted into a large BAC we have which contains the entire 3'end and extends all the way to DQ52—thus, this BAC contains all the control elements inside the expressing allele of 9921 except possible unknown elements 5' to DQ52. This newly modified large BAC could be transfected into the 9921 cell line and assayed for transcription efficiency compared to the endogenous level. In this case, the transcription of the μ heavy chain should be used instead of the α heavy chain in comparison to the endogenous γ2a transcription. If this experiment shows that the transgene made in this large BAC is expressed in a copy number dependent, position independent manner, further dissection of the roles of known elements (enhancers, MARs, etc) could be conducted.

4.4.2 BEμ expression did not reach the endogenous level.

We also observed that the expression levels of most Blxhs4 and BEμ transgenes were substantially lower than the endogenous γ2a transcript from 9921 (23.7% ± 20.7% for Blxhs4 and 30.3% ± 20.6%) on a per copy basis. The fact that the transgenes are not expressed at the same level as the expressing allele of 9921 is rather surprising because BACs B1-8α and Blxhs4 contain all the control elements with enhancer activity in their natural position and orientation as in the expressing allele of 9921, and BAC BEμ has an extra Eμ enhancer. There are two possibilities: one is that there are unknown elements

residing 5' to the 3'IgH regulatory region that are present in the endogenous locus of 9921 that play an important role in opening chromatin and insulating the IgH locus from the chromatin effect from adjacent chromatin.

Alternatively, the opening of the IgH locus could be a developmental process. Thus the fact that our transgene is not expressed as strongly as the endogenous Ig heavy chain gene could be simply due to the lack of the developmental process that confers the proper expression status on the transgene. As mentioned in the introduction, the IgH locus is opened in a stepwise manner. The full activation of a transgene might require a passage through earlier stages of B cell development. It was shown that a 155kb YAC containing the β-globin locus did not express β-globin in MEL cells before it experienced a passage through a non-erythroid cellular environment (Vassilopoulos et al., 1999). In the transgenic mice study using the 230 kb BAC transgene mentioned above, the serum level IgG1 after immunization reached that of the endogenous level (Dunnick et al., 2004). This result suggested that elements 3' to the D_H region are sufficient to drive transgene expression and class switch recombination at the endogenous level after going through development. Again, we need to take into the consideration that γ1 is independently regulated. Therefore, we must consider the possibility that the reason for the lowered expression level observed in the BAC transgenes is because the expressing allele of 9921 lost Eμ and other MARs within the constant region after the establishment of the expression of the transgene, while the BAC transgenes are integrated into integration sites to be expressed for the first time without the various preparation stages that the endogenous locus went through.

4.5 hs4 deletion did not greatly affect expression of the BAC Igα gene.

In this and previous studies, it has been demonstrated that the 3' IgH regulatory region is both necessary and sufficient to maintain transcription of endogenous IgH genes. It is also sufficient to establish transcription of transgenes in their own integration sites in late stage B cells. In cell lines bearing an IgH locus that had spontaneously lost Eμ, transcription of the IgH gene was preserved (Aguilera et al, 1985; Eckhardt and Birshtein 1985; Zaller and Eckhardt 1985). Furthermore, replacement of hs1,2 by the pgk-neo^r gene abolished transcription from the Ig secreting cell line 9921, suggesting that hs1,2 is required for high level transcription in Ig secreting cells. Alternatively, DNA sequences downstream of hs1,2, including the pair of enhancers hs3b and hs4, are required, and their transcription enhancement activity was diverted to transcribe the neo^r gene (Lieberson et al., 1995). In another experiment, the endogenous IgH locus of plasmactoma cell line LP1.2 lost 30 kb of DNA downstream of Cα, and the transcription level reduced to 1/8 to 1/4 the levels seen prior to this deletion (Gregor and Morrison 1985). It should be noted that the deletion in this special case extended to about 1.2kb downstream of hs4 as judged subsequently by sequence analysis, suggesting that the newly-discovered insulator encompassing HS5 through HS7 remained in the locus (Garrett et al., 2005). These observations demonstrated almost beyond doubt that the 3'IgH regulatory region was responsible for full activation of the expression of the endogenous IgH gene, unless, in a small chance, that hs4 itself or its synergy with downstream HS5~HS7 shields the 3'end of the locus, so that the decreased expression level resulted from negative effect of chromatin 3' to the IgH locus (See discussion below). In at least three cases, it has been shown that mini-cassettes containing three or

four known 3'IgH enhancers could guarantee the transcription of the transgenes linked to them (Madisen and Groudine 1994; Chauveau et al., 1999; Shi and Eckhardt 2001). Finally, in the current study, it was shown that the 3'IgH regulatory region in its natural configuration can drive transcription of the transgene in most integration sites in the plasmacytoma cell line 9921.

In search for the critical components of the 3'IgH regulatory region, transgenic mice studies have shown that neither hs3a nor hs1,2 alone is required for CSR (Manis et al, 1998), while the enhancer pair hs3b and hs4 is required for CSR to most heavy chain classes, probably due to lack of sterile transcription (Pinuad et al., 2001). Stable transfection assays in a plasmacytoma cell line demonstrated that the same enhancer pair is required for maintaining transcription of transgenes at their own integration sites (Shi and Eckhardt 2001).

In order to further dissect the critical components of this pair of enhancers, we replaced hs4 in the BAC B1-8α with floxed hs4. hs4 was chosen for its strong enhancer activity in transient transfection assays, its early display of DNaseI hypersensitivity, and the possible redundancy between hs3a and hs3b. To our surprise, hs4 deletion by CRE-loxP mediated deletion did not extinguish transcription of the transgene in any one of the integration sites. Rather four out five clones tested showed some reduction in transcription (20-50%), while one displayed a 40% increase in transcription level. Thus, hs4 is not required to maintain high level transcription from a transgene carried in a BAC in 9921, a cell line representing the last stage of B cell development.

Before we can draw further inferences, we need to take into consideration how comparable this experiment is to the previous experiments. (Shi and Eckhardt, 2001;

Pinaud et al., 2001). If we assume that the four 3'IgH enhancers carry all or most of the functions of the 3'IgH regulatory region, then the BAC transgene should be equivalent to the mini-construct used in Shi and Eckhardt (2001). An extrapolation from the result obtained from this previous experiment would be that the enhancer pair hs3b and hs4 is required for maintaining high level transcription from the BAC. Then we can reach an interpretation that either hs3b alone or the pair hs3b and hs4 is required for this process. Given the result obtained from hs4 deletion from the BAC, it can be postulated that either hs3b or the synergy among the three 3'IgH enhancers hs3a, hs1,2 and hs3b can compensate for the loss of hs4. It is unlikely that hs3b itself can carry out such function.

In vitro binding of transcription factors has been studied for mouse and human hs4 and for hs3b (Michaelson et al, 1996; Gordon et al., 2003; Sepulveda et al., 2004). As shown in Table I in the introduction, hs3b contains almost every transcription factor binding site that is found in hs4, except BSAP and NF-κb binding sites. BSAP is not expressed in plasma cells, ruling out the possibility that BSAP binding could distinguish the function of hs3b from hs4 in plasma cells, while NF-κb accounted for 50% of the activity of hs4 in transient transfection assays (Michaelson et al., 1996). YY1 was shown to bind hs3b *in vitro* and is responsible for activation of hs3b by LPS in transient transfection assays (Gordon et al., 2003). But currently, it is not known whether hs4 has YY1 binding sites and whether YY1 is required to maintain transcription in the IgH locus. In addition, it has been shown that the outcome of binding of a certain transcription factor to its cognate site can be influenced by the presence of other transcription factors (e.g. Singh and Birshtein 1996). The difference in the relative orders and distances between identical transcription factors in hs3b and hs4 could determine the difference in their

transcription enhancement activity. Enhancer activity of hs3b has indeed been shown to be much weaker than that of hs4 in transient transfection assay (Ong et al., 1998). In summary, there is no evidence at present that suggests that hs3b is critical for maintaining high level IgH transcription in plasma cells. Therefore, it is more likely that hs3b synergizes with other elements in the 3'IgH RR to compensate for the loss of hs4. Evidence exists to show that the synergy among weaker enhancers could compensate for the loss of a strong one. Each 3'IgH enhancer is a weak enhancer by transient transfection assay, but the combination of the four was equivalent in strength to Eμ (Ong et al., 1998). hs3a and hs3b seem to have a synergistic effect on transcription, although two copies of hs3 (equivalent to hs3b) behaved much stronger than the hs3a-hs3b combination (Michaelson et al., 1996). It would be interesting to find out whether hs3a and hs3b, in their natural 3'IgH configuration, can exert similar synergy.

It is possible that the 3'IgH RR in BAC B1-8α and Blxhs4 is not functionally equivalent to the cluster of the four 3'IgH enhancers in mini-locus constructs, and, therefore, it is possible that hs3b/hs4 deletion in this BAC would not result in the dramatic reduction in gene expression seen in earlier studies with a mini-locus (Shi and Eckhardt, 2001). First, the cluster of enhancers in the mini-locus lacks the putative insulator consisting of hs5 through hs7 and residing at the 3' end of the IgH locus, possibly rendering the mini-locus transgene more susceptible to position effect. Second, the 26kb highly repetitive sequence in the IgH locus was preserved in BAC Blxhs4 after hs4 deletion, and this may play a role in compensating for the loss of hs4, although the function of such highly repetitive sequences which might form a higher-order DNA structure has not been determined. A new BAC is being constructed to replace the

enhancer pair hs3b and hs4 with floxed hs3b and hs4. Transfection and double deletion analysis of this new BAC would determine whether the pair of enhancers hs3b and hs4 is required for IgH gene activity in the natural context and whether BAC B1-8α is equivalent to the mini-construct in terms of requirements for maintenance of transcription.

Single enhancer deletion of hs3a, or of hs1,2 did not yield any effect on class switch recombination (Manis et al., 1998). Clean double deletion of the pair of enhancers hs3b and hs4 dramatically reduced CSR to most classes of heavy chain (Pinaud et al., 2001). It would be interesting to determine whether hs3b or hs4 single deletions affect CSR.

Although hs4 possesses enhancer activity, its primary function in the endogenous locus might not be as a transcriptional enhancer for the V_H promoter. What is its physiological function, then, in the endogenous locus? The variation in transgene expression among the clones following deletion of hs4 was 32%, in contrast to 18% obtained before deletion of hs4. This increase in variability was mainly due to clone 6.1's increased transcription level, suggesting that hs4 might participate in insulating the locus. Although hs4 was shown not to be an enhancer blocker in the colony survival assay, the ability of hs4 to bind CTCF *in vitro* and the fact that deletion of the IgH RR hs4 compromised the ability of the chicken β-globin insulator 5'HS4 to insulate an IgH transgene suggest that the IgH RR hs4 might cooperate with other elements, such as IgH hs5~7 and the chicken β-globin insulator to shield IgH genes from the influence of adjacent chromatin (Garrett et al., 2005; Truffinet et al., 2005). The histone methylation pattern suddenly stopped at hs4 when extending from DNA sequences 3' to hs4 in pro-B and pre-B cells, which also supports the idea that hs4 could act like an insulator (Garrett

et al., 2005). In addition, the need for 2x chicken β-globin 5'HS4 insulator on each end of a transgene suggests the need for multiple and complex protein -protein interactions to completely insulate a transgene (Recillas-Targa et al., 2002).

Additional support for the notion that hs4 serves an insulator function comes from the finding that the insertion of a neo^r gene 2kb downstream of hs4 (between hs4 and the putative enhancer blocker) did not affect CSR while replacement of hs3a or hs1,2 with the neo^r gene abolished CSR, due to lack of sterile transcription (Manis et al., 2003). One of the arguments against hs4's serving as an insulator, however, is that there may be elements downstream of hs4 that contribute to the process of CSR. neo^r replacement of the hs3b/hs4 enhancers had a more dramatic effect on CSR than did clean deletion of these enhancers except in the case of switching to γ2b (Pinaud et al., 2001). This result suggested that neo^r blocked the interaction of the CSR machinery with unknown elements downstream of hs4 which were required for efficient CSR. If this were true, it would argue against an insulator role for hs4, since if hs4 were an insulator, it would prevent the interaction of I exon promoters with these downstream cis-elements. However, several lines of evidence suggest that elements located downstream of hs4 and required for CSR do not exist. neo^r insertion 2kb downstream of hs4 did not affect CSR, while replacement at hs4 drastically reduced CSR. If such postulated cis-elements exist, they must lie within 2kb downstream of hs4. However, extensive DNaseI hypersensitive assays did not reveal any DNaseI hypersensitive sites within 2kb downstream of hs4 (Garrett et al., 2005). In addition, in the *in vitro* LPS plus cytokine stimulation assay, efficiency of CSR to all classes was similar for the neo^r-replacement and clean-deletion mice, suggesting that the

dramatic difference between neo^r and clean deletion does not result from differences in the molecular events happening at the IgH locus (Pinaud et al., 2001).

4.6 Perspectives

4.6.1 Redundancy

So far, none of the 3'IgH enhancers has been proved to be indispensable for CSR or transcription in its natural context. It is not unprecedented that elements previously thought to be critical for certain cellular functions in transgenes turned out to be dispensable in a more natural context. The entire β-globin 5' LCR is required for opening up the chromatin structure in every transgene integration site. However, when the entire LCR was deleted from the endogenous locus, although the level of transcription decreased dramatically, such manifestations of chromatin configuration as histone acetylation and general sensitivity to DNaseI remained largely unchanged (Epner et al., 1998). Therefore, although the β-globin LCR is sufficient to confer transgene position-independent expression, some unknown elements in the locus might be redundant and could compensate the LCR for certain functions such as chromatin opening activity.

Redundancy could be a general phenomenon in regulating tissue specific genes. From an evolutionary point of view, mutations occurring in non-coding regions are easier tolerated than those occurring in coding regions. DNA sequences in individual enhancers are not well preserved from mouse to human (Sepulveda et al., 2005). For example, comparison between mouse and human hs4 showed that human hs4 lacks an octamer binding site and a BSAP binding site (Sepulveda et al., 2004). Rather, the repetitive feature of the DNA sequences seems to be conserved. On one hand, gains of function result from accumulated mutations that transform a non-protein binding site into a protein

binding site. The binding of a transcription factor to that site in a particular tissue is only possible when transcription factor is expressed in that tissue. Functional cis-elements are clusters of protein binding sites—only when multiple mutations have accumulated to a certain degree can a stretch of DNA sequence become a functional element. Therefore numerous putative protein binding sites could accumulate without devastating effect on organisms and species. The best tolerated newly-formed functional cis-element would be a redundant one. In a rare case, this element could gain evolutionary advantage for the species and remain in the genome. On the other hand, redundancy gains the species evolutionary advantage against loss of function. For example, if the chromatin opening element in the β-globin locus were mutated, redundant elements farther 5' could compensate for its loss, thus β-globin transcription would remain largely unchanged, and the organism would survive.

Therefore, one should not be surprised to see tremendous redundancy within the IgH locus. It is possible that even double deletion of hs3b and hs4 in the BAC system would not yield any effect on IgH transcription. In this case, systematic mapping of the regions necessary for the gene's activity (that is, essential components of the partial LCR) could be conducted by replacing cis-elements in various positions of the 3'IgH RR with loxP flanked cis-elements. Following CRE-loxP mediated recombination, single cis-elements or large pieces of DNA between loxP flanked elements could be deleted. By this method, we will be able to collect definitive information on the redundancy and functional synergy among different regions, and even be able to explore the possible role played by higher DNA structures.

4.6.2 Division of Labor

Studies of several LCRs such as the β-globin LCR and T cell receptor alpha (TCRα) LCR revealed clear divisions of labor among different DNaseI hypersensitive sites within the LCRs. The enhancer activity of the β-globin locus was attributed to 5'HS2, 3, and 4. 5'HS5 is a chromatin insulator, and the function of 5'HS1 remains unknown (Li et al., 2002). In the TCR locus, while the 3' HS2 to HS6 contain non-specific chromatin opening activity, the 5' end of the LCR region is responsible for cell-type specific chromatin opening activity (Ortiz et al., 1997). Even within a single hypersensitive site, such as the chicken β-globin 5'HS4, the insulator function is divided into enhancer blocking and chromatin barrier activities (Recillars-Targa, et al., 2002).

In the IgH locus, Eμ is required for efficient V to DJ recombination (Chen et al., 1993; Serwe and Sablitzky, 1993). The pair of enhancers hs3b and hs4 is required for Class Switch Recombination (Pinaud et al., 2001). However, only with the recent finding of a putative insulator, have we just begun to appreciate a division of labor at the molecular level (Garrett et al., 2005). In this regard, the studies on the IgH locus seem to lag behind those on the β-globin LCR and TCRα LCR, probably due to the large size and the complexity of various recombination, somatic hypermutation, and transcription events that take place in the IgH locus. On the other hand, IgH LCR might be different from these other LCRs in that functional synergy and/or higher DNA structure might play more important roles than individual cis-elements in the regulation of these events. Bacterial Artificial chromosomes (BAC) not only provide a tool to carry the large 3'IgH regulatory region but also can be labeled with DNA tags that enable us to distinguish the BAC transgenes from the endogenous IgH genes to study changes in chromatin structures

in various places. It also can serve as a tool for us to manipulate the higher DNA structure by changing long DNA sequences with homologous recombination in bacteria. In the near future, it can be foreseen that not only mechanisms reminiscent of those involved in β-globin and TCRα LCR functions but also those distinct in the regulation of IgH genes will be discovered.

References:

Aguilera RJ, Hope TJ, Sakano H. (1985). Characterization of immunoglobulin enhancer deletions in murine plasmacytomas. EMBO J. *4*, 3689-93.

Andersson, T., Furebring, C., Borrebaeck, C., Pettersson, S. (1999) Temporal expression of a V_H promoter-Cμ transgene linked to the IgH HS1,2 enhancer. Mol Immunol. *36*, 19-29.

Artandi SE, Cooper C, Shrivastava A, and Calame K (1994). The basic helix-loop-helix-zipper domain of TFE3 mediates enhancer-promoter interaction. Mol Cell Biol. *14*, 7704–16.

Arulampalam V, Eckhardt L, and Pettersson S. (1994a). The enhancer shift: a model to explain the developmental control of IgH gene expression in B-lineage cells. Immunol Today. *18*, 549-54.

Arulampalam V, Grant PA, Samuelsson A, Lendahl U, and Pettersson S. (1994b). Lipopolysaccharide-dependent transactivation of the temporally regulated immunoglobulin heavy chain 3' enhancer. Eur J Immunol. *24*, 1671-7.

Banerji J, Olson L, Schaffner W. (1983) A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell. *33*: 729-40.

Barton MC and Emerson BM. (1994). Regulated expression of the beta-globin gene locus in synthetic nuclei. Genes Dev. *8*, 2453-65.

Bender MA, Mehaffey MG, Telling A, Hug B, Ley TJ, Groudine M, Fiering S. (2000). Independent formation of DnaseI hypersensitive sites in the murine beta-globin locus control region. Blood. *95*, 3600-4.

Brodeur PH, Osman GE, Mackle JJ, Lalor TM. (1988). The organization of the mouse Igh-V locus. Dispersion, interspersion, and the evolution of VH gene family clusters. J Exp Med. *168*, 2261-78.

Bungert J, Tanimoto K, Patel S, Liu Q, Fear M, Engel JD. (1999). Hypersensitive site 2 specifies a unique function within the human beta-globin locus control region to stimulate globin gene transcription. Mol Cell Biol. *19*, 3062-72

Burhans WC, Vassilev LT, Wu J, Sogo JM, Nallaseth FS, DePamphilis ML.(1991). Emetine allows identification of origins of mammalian DNA replication by imbalanced DNA synthesis, not through conservative nucleosome segregation. EMBO J. *10,* 4351-60.

Cazzaniga G, Daniotti M, Tosi S, Giudici G, Aloisi A, Pogliani E, Kearney L, Biondi A. (2001).The paired box domain gene PAX5 is fused to ETV6/TEL in an acute lymphoblastic leukemia case. Cancer Res. *61*, 4666-70.

Chauveau, C., and Cogne, M. (1996). Palindromic structure of the IgH 3' locus control region. Nature Genet. *14*, 15-6.

Chauveau, C., Jansson, E. A., Muller, S., Cogne, M., and Pettersson, S. (1999). Cuttin Edge: Ig heavy chain 3' HS1-4 directs correct spatial and position-independent expression of a linked transgene to B lineage cells. J. Immunol. *163*, 4637-4641.

Chen, J., Young, F., Bottaro, A., Stewart, V., Smith, R. K., and Alt, F. W. (1993). Mutations of the intronic IgH enhancer and its flanking sequences differentially affect accessibility of the JH locus. Embo J *12*, 4635-45.

Chipev CC and Wolffe AP. (1992). Chromosomal organization of Xenopus laevis oocyte and somatic 5S rRNA genes *in vivo*. Mol Cell Biol. *12*:45-55.

Chowdhury D and Sen R. (2001). Stepwise activation of the immunoglobulin mu heavy chain gene locus. EMBO J. *20*, 6394-403.

Cockerill PN. (1990). Nuclear matrix attachment occurs in several regions of the IgH locus. Nucleic Acids Res. *18*, 2643-8

Cogne, M., Lansford, R., Bottaro, A., Zhang, J., Gorman, J., Young, F., Cheng, H. L., and Alt, F. W. (1994). A class switch control region at the 3' end of the immunoglobulin heavy chain locus. Cell *77*, 737-47.

Dang CV. (1999). *c-myc* target genes involved in cell growth, apoptosis, and metabolism. Mol Cell Biol. 1999 *19*, 1-11.

Darby MK, Andrews MT, Brown DD. (1988). Transcription complexes that program Xenopus 5S RNA genes are stable *in vivo*. Proc Natl Acad Sci U S A. 85, 5516-20.

Dariavach, P., Williams, G. T., Campbell, K., Pettersson, S., and Neuberger, M. S. (1991). The mouse IgH 3'-enhancer. Eur J Immunol *21*, 1499-504.

DeKoter RP, Lee HJ, Singh H. (2002). PU.1 regulates expression of the interleukin-7 receptor in lymphoid progenitors. Immunity. *16*, 297-309.

Delaval K and Feil R.(2004). Epigenetic regulation of mammalian genomic imprinting. Curr Opin Genet Dev. *14*, 188-95.

Dildrop R, Ma A, Zimmerman K, Hsu E, Tesfaye A, DePinho R, Alt FW. (1989). IgH enhancer-mediated deregulation of N-myc gene expression in transgenic mice: generation of lymphoid neoplasias that lack *c-myc* expression. EMBO J. *8*: 1121-8.

Dillon, N., Trimborn, T., Strouboulis, J., Fraser, P., and Grosveld, F. (1997). The effect of distance on long-range chromatin interactions. Mol Cell *1*, 131-9.

Driscoll MC, Dobkin CS, Alter BP. (1989).Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites. Proc Natl Acad Sci U S A. *86*, 7470-4.

Dunaway M, and Droge P. (1989). Transactivation of the Xenopus rRNA gene promoter by its enhancer. Nature. *341*, 657-9.

Dunnick WA, Shi J, Graves KA, Collins JT. (2004). Germline transcription and switch recombination of a transgene containing the entire H chain constant region locus: effect of a mutation in a STAT6 binding site in the gamma 1 promoter. J Immunol. *173*, 5531-9.

Eckhardt, L. A. (1992). Immunoglobulin gene expression only in the right cells at the right time. FASEB J *6*, 2553-2560.

Eckhardt, L. A., and Birshtein, B. K. (1985). Independent immunoglobulin class-switch events occurring in a single myeloma cell line. Mol.Cell. Biol. *5*, 856-868.

Edelman GM, Cunningham BA, Gall WE, Gottlieb PD, Rutishauser U, Waxdal MJ. (1969). The covalent structure of an entire gammaG immunoglobulin molecule. Proc Natl Acad Sci U S A. *63*, 78-85.

Ellis, J., K. C. Tan-Un, A. Harper, D. Michalovich, N. Yannoutsos, S. Philipsen, F. Grosveld. 1996. A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human {beta}-globin locus control region. EMBO J. *15*, 562

Engel JD, Tanimoto K. (2000). Looping, linking, and chromatin activity: new insights into beta-globin locus regulation. Cell. *100*, 499-502.

Epner E, Reik A, Cimbora D, Telling A, Bender MA, Fiering S, Enver T, Martin DI, Kennedy M, Keller G, Groudine M. (1998). The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. Mol Cell. *2*, 447-55.

Ernst, P. and Smale, S. T. (1995). Combinatorial regulation of transcription II: the immunoglobulin mu heavy chain gene. Immunity *2*, 427-438.

Fernandez LA, Winkler M, Grosschedl R. (2001). Matrix attachment region-dependent function of the immunoglobulin mu enhancer involves histone acetylation at a distance without changes in enhancer occupancy. Mol Cell Biol. *21*:196-208

Festenstein, R., Tolaini, M., Corbella, P., Mamalaki, C., Parrington, J., Fox, M., Miliou, A., Jones, M., and Kioussis, D. (1996). Locus control region function and heterochromatin-induced position effect variegation. Science *271*, 1123-5.

Flores O, Lu H, Reinberg D. (1992). Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIH. J Biol Chem. *267*, 2786-93.

Forrester WC, van Genderen C, Jenuwein T, Grosschedl R.(1994). Dependence of enhancer-mediated transcription of the immunoglobulin mu gene on nuclear matrix attachment regions. Science. *265*, 1221-5.

Garrett FE, Emelyanov AV, Sepulveda MA, Flanagan P, Volpi S, Li F, Loukinov D, Eckhardt LA, Lobanenkov VV, Birshtein BK. (2005). Chromatin architecture near a potential 3' end of the igh locus involves modular regulation of histone modifications during B-Cell development and *in vivo* occupancy at CTCF sites. Mol Cell Biol. *25*, 1511-25.

Giannini, S. L., Singh, M., Calvo, C. F., Ding, G., and Birshtein, B. K. (1993). DNA regions flanking the mouse Ig 3' alpha enhancer are differentially methylated and DNAase I hypersensitive during B cell differentiation. J Immunol *150*, 1772-80.

Giglioni, B., Casini, C., Mantovani, R., Merli, S., Comi, P., Ottolenghi, S., Saglio, G., Camaschella, C., and Mazza, U. (1984). A molecular study of a family with Greek hereditary persistence of fetal hemoglobin and β-thalassemia. EMBO J. *3*, 2641-2645.

Gillies SD, Morrison SL, Oi VT, Tonegawa S. (1983). A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell. *33*, 717-28.

Gordon SJ, Saleque S, Birshtein BK. (2003). Yin Yang 1 is a lipopolysaccharide-inducible activator of the murine 3' Igh enhancer, hs3. J Immunol. *170*, 5549-57.

Gregor, P. D., and Morrison, S. L. (1986). Myeloma mutant with a novel 3' flanking region: loss of normal sequence and insertion of repetitive elements leads to decreased transcription but normal processing of the alpha heavy-chain gene products. Mol Cell Biol *6*, 1903-1916.

Gregory PD and Horz W. (1998). Chromatin and transcription--how transcription factors battle with a repressive chromatin environment. Eur J Biochem. *251*, 9-18.
Gross DS, and Garrard WT. (1988). Nuclease hypersensitive sites in chromatin.
Annu Rev Biochem. *57*, 159-97.

Grosschedl R, and Marx M. (1988). Stable propagation of the active transcriptional state of an immunoglobulin mu gene requires continuous enhancer function. Cell. *55*, 645-54

Grosveld F, van Assendelft GB, Greaves DR, Kollias G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell. *51*, 975-85.

Guglielmi L, Le Bert M, Cogne M, Denizot Y. (2003a). Effect of the Emu IgH enhancer on expression of a GFP reporter gene in transfected B cells and transgenic mice. Immunol Lett *86,* 77-83.

Guglielmi L, Le Bert M, Comte I, Dessain ML, Drouet M, Ayer-Le Lievre C, Cogne M, Denizot Y. (2003b). Combination of 3' and 5' IgH regulatory elements mimics the B-specific endogenous expression pattern of IgH genes from pro-B cells to mature B cells in a transgenic mouse model. Biochim Biophys Acta. *1642*, 181-90.

Guglielmi L, Truffinet V, Cogne M, and Denizot Y. (2003c) The beta-globin HS4 insulator confers copy-number dependent expression of IgH regulatory elements in stable B cell transfectants. Immunol Lett. *89*:119-23.

Ha I, Roberts S, Maldonado E, Sun X, Kim LU, Green M, and Reinberg D. (1993). Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. Genes Dev. *7*, 1021-32.

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J.M., and Tilghman S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature *405*, 486-489.

Herr W, Sturm RA, Clerc RG, Corcoran LM, Baltimore D, Sharp PA, Ingraham HA, Rosenfeld MG, Finney M, Ruvkun G, et al. (1988). The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and Caenorhabditis elegans unc-86 gene products. Genes Dev. *2*, 1513-6.

Hochheimer A and Tjian R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev. *17*, 1309-20.

Honjo, T. and Matsuda (1995) Immunoglobulin heavy chain loci of mouse and human. Immunoglobulin Genes 2nd Edition, Academic Press Limited, (edited by T. Honjo, and F W. Alt) 145-172.

Hu JS, Olson EN, Kingston RE. (1992) HEB, a helix-loop-helix protein related to E2A and ITF2 that can modulate the DNA-binding ability of myogenic regulatory factors. Mol Cell Biol. *12*, 1031-42.

Ishida N, Ueda S, Hayashida H, Miyata T, Honjo T. (1982)The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison with the human epsilon gene sequence. EMBO J. *1*,1117-23

Jenuwein T, Forrester WC, Qiu RG, Grosschedl R. (1993).The immunoglobulin mu enhancer core establishes local factor access in nuclear chromatin independent of transcriptional stimulation.Genes Dev. *7*, 2016-32.

Jenuwein T, Forrester WC, Fernandez-Herrero LA, Laible G, Dull M, Grosschedl R. (1997). Extension of chromatin accessibility by nuclear matrix attachment regions. Nature. 1997 *385*, 269-72.

Jenuwein T and Allis CD. (2001). Translating the histone code. Science. *293*, 1074-80.

Jung S, Rajewsky K, Radbruch A. (1993). Shutdown of class switch recombination by deletion of a switch region control element. Science. *259*, 984-7

Kadesch T, Zervos P, Ruezinsky D. (1986). Functional analysis of the murine IgH enhancer: evidence for negative control of cell-type specificity. Nucleic Acids Res. *14*, 8209-21.

Kerppola TK, Curran T. (1991) Fos-Jun heterodimers and Jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. Cell. *66*,317-26.

Kim, U., Qin, X. F., Gong, S., S., S., Luo, Y., Nussenzweig, M., and Roeder, R. G. (1996). The B-cell-specific transcriptional coactivator OCA-B/ OBF-1/ Bob-1 is essential for normal production of immunoglobulin isotypes. Nature *383*, 542-547.

Kofler R, Geley S, Kofler H, Helmberg A. (1992). Mouse variable-region gene families: complexity, polymorphism and use in non-autoimmune responses. Immunol Rev. *128*, 5-21.

Langdon WY, Harris AW, Cory S, Adams JM. (1986) The *c-myc* oncogene perturbs B lymphocyte development in E-mu-myc transgenic mice. Cell. *47*, 11-8.

Laurencikiene J, Deveikaite V, Severinson E. (2001). HS1,2 enhancer regulation of germline epsilon and gamma2b promoters in murine B lymphocytes: evidence for specific promoter-enhancer interactions. J Immunol. *167*, 3257-65.

Lenardo M, Pierce JW, Baltimore D. (1987). Protein-binding sites in Ig gene enhancers determine transcriptional activity and inducibility. Science. *236*, 1573-7

Li Q, Verma IM. (2002a). NF-kappaB regulation in the immune system. Nat Rev Immunol. 2, 725-34.

Li Q, Peterson KR, Fang X, Stamatoyannopoulos G. (2002b). Locus control regions. Blood. *100*, 3077-86.

Lieberson, R., Giannini, S. L., Birshtein, B. K., and Eckhardt, L. A. (1991). An enhancer at the 3' end of the mouse immunoglobulin heavy chain locus. Nucleic Acids Res *19*, 933-7.

Lieberson, R., Ong, J., Shi, X., and Eckhardt, L. A. (1995). Immunoglobulin gene transcription ceases upon deletion of a distant enhancer. Embo J *14*, 6229-38.

Linderson Y, Eberhard D, Malin S, Johansson A, Busslinger M, Pettersson S. (2004). Corecruitment of the Grg4 repressor by PU.1 is critical for Pax5-mediated repression of B-cell-specific genes. EMBO J 5, 291-6.

Lloyd A and Sakonju S. (1991). Characterization of two Drosophila POU domain genes, related to oct-1 and oct-2, and the regulation of their expression patterns. Mech Dev. *36*, 87-102

Lorenz M, Jung S, Radbruch A. (1995). Switch transcripts in immunoglobulin class switching. Science. *267*,1825-8.

Lorenz M and Radbruch A. (1996).Developmental and molecular regulation of immunoglobulin class switch recombination. Curr Top Microbiol Immunol. *217*, 151-69.

Luo, Y., Fujii, H., Gerster, T., and Roeder, R. G. (1992). A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. Cell *71*, 231-241.

Luo, Y., and Roeder, R. G. (1995). Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator OCA-B. Mol Cell Biol *15*, 4115-4124.

Lutzker S, Alt FW. (1988). Structure and expression of germ line immunoglobulin gamma 2b transcripts. Mol Cell Biol. *8*, 1849-52.

Madisen, L., and Groudine, M. (1994). Identification of a locus control region in the immunoglobulin heavy- chain locus that deregulates *c-myc* expression in plasmacytoma and Burkitt's lymphoma cells. Genes Dev *8*, 2212-26.

Madisen L, Krumm A, Hebbes TR, Groudine M. (1998). The immunoglobulin heavy chain locus control region increases histone acetylation along linked *c-myc* genes. Mol Cell Biol. *18*, 6281-92.

Manis, J. P., van der Stoep, N., Tian, M., Ferrini, R., Davidson, L., Bottaro, A., and Alt, F. W. (1998). Class switching in B cells lacking 3' immunoglobulin heavy chain enhancers. J Exp Med *188*, 1421-31.

Manis JP, Michaelson JS, Birshtein BK, and Alt FW. (2003). Elucidation of a downstream boundary of the 3' IgH regulatory region. Mol Immunol. *39*, 753-60.

Marecki S, McCarthy KM, Nikolajczyk BS. (2004). PU.1 as a chromatin accessibility factor for immunoglobulin genes. Mol Immunol. *40*, 723-31.

Matthias, P., and Baltimore, D. (1993). The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the alpha constant region. Mol. Cell. Biol. *13*, 1547-1553.

Michaelson, J. S., Giannini, S. L., and Birshtein, B. K. (1995). Identification of 3' alpha-hs4, a novel Ig heavy chain enhancer element regulated at multiple stages of B cell differentiation. Nucleic Acids Res *23*, 975-81.

Michaelson, J. S., Singh, M., Snapper, C. M., Sha, W. C., Baltimore, D., and Birshtein, B. K. (1996). Regulation of 3' IgH enhancers by a common set of factors, including kappa B-binding proteins. J Immunol *156*, 2828-39.

Mills, F. C., Harindranath, N., Mitchell, M., and Max, E. E. (1997). Enhancer complexes located downstream of both human immunoglobulin Calpha genes. J Exp Med *186*, 845-58.

Milot E, Fraser P, Grosveld F. (1996) Position effects and genetic disease. Trends Genet. *12*, 123-6.

Misulovin Z, Yang XW, Yu W, Heintz N, Meffre E. A rapid method for targeted modification and screening of recombinant bacterial artificial chromosome. J Immunol Methods. *257*, 99-105.

Monaco AP, Larin Z. (1994). YACs, BACs, PACs and MACs: artificial chromosomes as research tools. Trends Biotechnol. 1994 *12*, 280-6.

Morcillo P, Rosen C, Baylies MK, Dorsett D. (1997). Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in Drosophila. Genes Dev. *11*, 2729-40.

Moroy T, Fisher P, Guidos C, Ma A, Zimmerman K, Tesfaye A, DePinho R, Weissman I, Alt FW. (1990). IgH enhancer deregulated expression of L-myc: abnormal T lymphocyte development and T cell lymphomagenesis. EMBO J. *9*, 3659-66.

Mostoslavsky R, Alt FW, Rajewsky K. (2004). The lingering enigma of the allelic exclusion mechanism. Cell. *118*, 539-44.

Mueller-Storm HP, Sogo JM, Schaffner W. (1989). An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. Cell. *58*, 767-77.

Muto A, Hoshino H, Madisen L, Yanai N, Obinata M, Karasuyama H, Hayashi N, Nakauchi H, Yamamoto M, Groudine M, Igarashi K. (1998). Identification of Bach2 as a

B-cell-specific partner for small maf proteins that negatively regulate the immunoglobulin heavy chain gene 3' enhancer. EMBO J. *17*, 5734-43.

Nelsen B, Tian G, Erman B, Gregoire J, Maki R, Graves B, Sen R. (1993). Regulation of lymphoid-specific immunoglobulin mu heavy chain gene enhancer by ETS-domain proteins. Science. *261*, 82-6.

Neuberger MS. (1983). Expression and regulation of immunoglobulin heavy chain gene transfected into lymphoid cells. EMBO J. *2*,1373-8.

Neuberger MS, Di Noia JM, Beale RC, Williams GT, Yang Z, Rada C. (2005). Somatic hypermutation at A.T pairs: polymerase error versus dUTP incorporation. Nat Rev Immunol. *5*, 171-8.

Nielsen, P. J., Georgiev, O., Lorenz, B., and Schaffner, W. (1996). B lymphocytes are impaired in mice lacking the transcriptional co- activator Bob1/OCA-B/OBF1. Eur J Immunol *26*, 3214-8.

Nikolajczyk BS, Sanchez JA, Sen R. (1999) ETS protein-dependent accessibility changes at the immunoglobulin mu heavy chain enhancer. Immunity. *11*, 11-20.

Nitschke L, Kestler J, Tallone T, Pelkonen S, Pelkonen J. (2001). Deletion of the DQ52 element within the Ig heavy chain locus leads to a selective reduction in VDJ recombination and altered D gene usage. J Immunol. *166*, 2540-52.

Nolan GP, Baltimore D. (1992). The inhibitory ankyrin and activator Rel proteins. Curr Opin Genet Dev. 2, 211-20.

Nye JA, Petersen JM, Gunther CV, Jonsen MD, Graves BJ. (1992). Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. Genes Dev. *6*, 975-90.

Oancea, A. E., Berru, M., and Shulman, M. J. (1997). Expression of the (recombinant) endogenous immunoglobulin heavy-chain locus requires the intronic matrix attachment regions. Mol Cell Biol *17*, 2658-68.

O'Connor M, Peifer M, Bender W. (1989). Construction of large DNA segments in Escherichia coli. Science. 244, 1307-12.

Ong, J., Stevens, S., Roeder, R. G., and Eckhardt, L. A. (1998). 3' IgH enhancer elements shift synergistic interactions during B cell development. J Immunol *160*, 4896-903.

Ortiz BD, Cado D, Chen V, Diaz PW, Winoto A. (1997). Adjacent DNA elements dominantly restrict the ubiquitous activity of a novel chromatin-opening region to specific tissues. EMBO J. *16*, 5037-45.

Pawlitzky, I Stanton, M. L. and Brodeur, P. H 2004 Clinical and Investigative Medicine 207, 490.

Peterson KR, Clegg CH, Navas PA, Norton EJ, Kimbrough TG, Stamatoyannopoulos G. (1996). Effect of deletion of 5'HS3 or 5'HS2 of the human beta-globin locus control region on the developmental regulation of globin gene expression in beta-globin locus yeast artificial chromosome transgenic mice. Proc Natl Acad Sci U S A. *93*, 6605-9.

Peterson MG, Tjian R. (1992). Transcription. The tell-tail trigger. Nature. 358, 620-1.
Pettersson, S., Cook, G. P., Bruggemann, M., Williams, G. T., and Neuberger, M. S. (1990). A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus. Nature *344*, 165-8.

Pettersson S, Cook GP, Bruggemann M, Williams GT, Neuberger MS. (1990). A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus. Nature. 344, 165-8.

Pinaud E, Khamlichi AA, Le Morvan C, Drouet M, Nalesso V, Le Bert M, Cogne M. (2001). Localization of the 3' IgH locus elements that effect long-distance regulation of class switch recombination. Immunity. *15*, 187-99.

Porton B, Zaller DM, Lieberson R, Eckhardt LA. (1990). Immunoglobulin heavy-chain enhancer is required to maintain transfected gamma 2A gene expression in a pre-B-cell line. Mol Cell Biol. *10*,1076-83.

Preud'homme JL, Petit I, Barra A, Morel F, Lecron JC, Lelievre E. (2000) Structural and functional properties of membrane and secreted IgD. Mol Immunol.37, 871-87.

Price DH, Sluder AE, Greenleaf AL. (1989). Dynamic interaction between a Drosophila transcription factor and RNA polymerase II. Mol Cell Biol. *9*, 1465-75.

Pugh BF. and Tjian R. (1990). Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell. *61*, 1187-97.

Qin, X. F., Reichlin, A., Luo, Y., Roeder, R. G., and Nussenzweig, M. C. (1998). OCA-B integrates B cell antigen receptor-, CD40L- and IL 4-mediated signals for the germinal center pathway of B cell development. Embo J *17*, 5066-75.

Rack KA, Salomon-Nguyen F, Radford-Weiss I, Gil MO, Schmitt C, Belanger C, Nusbaum S, Vekemans M, Valensi F, Macintyre EA. (1998). FISH detection of chromosome 14q32/IgH translocations: evaluation in follicular lymphoma. Br J Haematol. 103, 495-504.

Radomska HS, Shen CP, Kadesch T, Eckhardt LA. (1994). Constitutively expressed Oct-2 prevents immunoglobulin gene silencing in myeloma x T cell hybrids.
Immunity. *1*, 623-34.

Recillas-Targa F, Bell AC, Felsenfeld G. (1999). Positional enhancer-blocking activity of the chicken beta-globin insulator in transiently transfected cells. Proc Natl Acad Sci *96*, 14354-9.

Recillas-Targa F, Pikaart MJ, Burgess-Beusse B, Bell AC, Litt MD, West AG, Gaszner M, Felsenfeld G. (2002). Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. Proc Natl Acad Sci U S A. *99*, 6883-8.

Reth, M., Hammerling, G. J., and Rajewsky, K. (1978). Analysis of the repertoire of anti-NP antibodiesin C57BL/6 mice by cell fusion. Characterization of antibody families in the primary and hyperimmune response. Eur. J. Immunol. *8*, 393-400

Ronai D, Berru M, Shulman MJ. (1999). Variegated expression of the endogenous immunoglobulin heavy-chain gene in the absence of the intronic locus control region. Mol Cell Biol. *19*, 7031-40

Rosenbaum H, Webb E, Adams JM, Cory S, Harris AW. (1989). N-myc transgene promotes B lymphoid proliferation, elicits lymphomas and reveals cross-regulation with *c-myc*. EMBO J. *8*, 749-55.

Saleque, S., Singh, M., Little, R. D., Giannini, S. L., Michaelson, J. S., and Birshtein, B. K. (1997). Dyad symmetry within the mouse 3' IgH regulatory region includes two virtually identical enhancers (C alpha3'E and hs3). J Immunol *158*, 4780-7.

Saleque, S., Singh, M., and Birshtein, B. K. (1999). Ig heavy chain expression and class switching *in vitro* from an allele lacking the 3' enhancers DNase I-hypersensitive hs3A and hs1,2. J Immunol *162*, 2791-803.

Sato S, Hasegawa M, Fujimoto M, Tedder TF, and Takehara K. 2000. Quantitative genetic variation in CD19 expression correlates with autoimmunity. J. Immunol. *165,* 6635-43.

Schebesta M, Heavey B, Busslinger M. (2002). Transcriptional control of B-cell development. Curr Opin Immunol. *14*, 216-23

Schubart, D. B., Rolink, A., Kosco-Vilbois, M. H., Botteri, F., and Matthias, P. (1996). B-cell-specific coactivator OBF-1/ OCA-B/ Bob-1 required for immune response and germinal centre formation. Nature *383*, 538-542.

Schubeler D, Groudine M, Bender MA. (2001). The murine beta-globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. Proc Natl Acad Sci U S A.*98*, 11432-7.

Scott EW, Simon MC, Anastasi J, Singh H. (1994). Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. Science. *265*, 1573-7.

Sen, R.,and Baltimore, D. (1989). Factors regulating immunoglobulin-gene transcription. In: *Immunoglobulin Genes* (Honjo, T., Alt, F.W., and Rabbitts, T.H., eds) pp.327-342, Academic, New York

Sepulveda MA, Emelyanov AV, Birshtein BK. (2004). NF-kappa B and Oct-2 synergize to activate the human 3' Igh hs4 enhancer in B cells. J Immunol. *172*, 1054-64.

Sepulveda MA, Garrett FE, Price-Whelan A, Birshtein BK. (2005) Comparative analysis of human and mouse 3' Igh regulatory regions identifies distinctive structural features. Mol Immunol. *42*, 605-15.

Serwe, M., and Sablitzky, F. (1993). V(D)J recombination in B cells is impaired but not blocked by targeted deletion of the immunoglobulin heavy chain intron enhancer. EMBO J *12*, 2321-7.

Sharif MN, Radomska HS, Miller DM, Eckhardt LA. (2001). Unique function for carboxyl-terminal domain of Oct-2 in Ig-secreting cells. J Immunol. *167*, 4421-9.

Shi X, Eckhardt LA. (2001). Deletional analyses reveal an essential role for the hs3b/hs4 IgH 3' enhancer pair in an Ig-secreting but not an earlier-stage B cell line. Int Immunol. 13, 1003-12.

Singh M, Birshtein BK. (1996). Concerted repression of an immunoglobulin heavy-chain enhancer, 3' alpha E(hs1,2). Proc Natl Acad Sci U S A. 93, 4392-7.

Snapper, C. M., Zalozowski, P., Rosas, F. R., Kehry, M. R., Tian, M., Baltimore, D., and Sha, W. C. (1996). B cells from p50/NF-kB knock-out mice have selective defects in proliferation, differentiation, germ line CH transcription, and Ig class-switching. J Immunol *156*, 183-191.

Sonoda E, Pewzner-Jung Y, Schwers S, Taki S, Jung S, Eilat D, Rajewsky K. (1997). B cell development under the condition of allelic inclusion. Immunity. *6*, 225-33.

Tang, H., and Sharp, P. A. (1999). Transcriptional regulation of the murine 3' IgH enhancer by OCT-2. Immunity *11*, 517-26.

Tanimoto K, Liu Q, Bungert J, Engel JD. (1999). Effects of altered gene order or orientation of the locus control region on human beta-globin gene expression in mice. Nature. *398*, 344-8.

Truffinet V, Guglielmi L, Cogne M, Denizot Y. (2005).The chicken beta-globin HS4 insulator is not a silver bullet to obtain copy-number dependent expression of transgenes in stable B cell transfectants. Immunol Lett. *96*, 303-4.

Vassilopoulos G, Navas PA, Skarpidi E, Peterson KR, Lowrey CH, Papayannopoulou T, Stamatoyannopoulos G. (1999). Correct function of the locus control region may require passage through a nonerythroid cellular environment. Blood.1999 *93*, 703-12

Wabl, M. R., and Burrows, P. D. (1984). Expression of immunoglobulin heavy chain at a high level in the absence of a proposed immunoglobulin enhancer element in cis. Proc. Natl. Acad.Sci. USA *81*, 2452-2455.

Willerford DM, Swat W, Alt FW. (1996) Developmental regulation of V(D)J recombination and lymphocyte differentiation. Curr Opin Genet Dev. *6*, 603-9.

Wolffe AP. (1993). Replication timing and Xenopus 5S RNA gene transcription *in vitro*. Dev Biol. *157*, 224-31.

Wood WI, Felsenfeld G. (1982). Chromatin structure of the chicken beta-globin gene region. Sensitivity to DNase I, micrococcal nuclease, and DNase II. J Biol Chem. *257,* 7730-6.

Yang XW, Model P, Heintz N. (1997). Homologous recombination based modification in Escherichia coli and germline transmission in transgenic mice of a bacterial artificial chromosome. Nat Biotechnol. *15*, 859-65.

Yu W, Nagaoka H, Jankovic M, Misulovin Z, Suh H, Rolink A, Melchers F, Meffre E, Nussenzweig MC. (1999). Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. Nature. *400*, 682-7.

Zaller, D. M., and Eckhardt, L. A. (1985). Deletion of a B-cell-specific enhancer affects transfected, but not endogenous, immunoglobulin heavy-chain gene expression. Proc. Nat. Acad. Sci. USA *82*, 5088-5092.

Zhang, J., Alt, F.W. and Honjo, T. (1995) Regulation of class switch recombination of the immunoglobulin heavy chain genes. Immunoglobulin Genes 2nd Edition, Academic Press Limited, (edited by T. Honjo, T W. Alt and T. H. Rabbitts.) 235-266.

Zhou J, Ermakova OV, Riblet R, Birshtein BK, Schildkraut CL. (2002). Replication and subnuclear location dynamics of the immunoglobulin heavy-chain locus in B-lineage cells. Mol Cell Biol. *22*, 4876-89.