The Dissertation Committee for Yonghwan Kim Certifies that this is the approved version of the following dissertation:

Identification and Characterization of a Positive Control Region for Activation Induced Cytidine Deaminase Mediated Gene Conversion in Chicken B Cells

Committee:

Ming Tian, Supervisor

Philip W. Tucker

Tanya T. Paull

Vishwanath Iyer

Y. Whitney Yin

Identification and Characterization of a Positive Regulatory Region for Activation Induced Cytidine Deaminase Mediated Gene Conversion in Chicken B Cells

by

Yonghwan Kim, B.S., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2009

Dedication

to my wife, Yumi

Acknowledgements

My God has been always there whenever I need Him. Later, I realized that He knows everything about me, but answers me when I call Him. I thank God for His great and intimate love for me and my family.

There are many whose support, encouragement, guidance and example make my dissertation possible. On the top of this list must be Ming Tian. Not only did he teach me how to read, think and do my research from beginning to end, but he also showed himself as a role model of good scientists. I have felt safe as he was always besides me and gave me a clue to move on to next level when I was struggle with problems. I appreciate what he has done for me and I will not forget that.

I would like to thank my committee members, Drs. Philip Tucker, Tanya Paull, Vishwanath Iyer and Whitney Yin for their invaluable comments, suggestion and encouragement during the processes. I am personally indebted to Philip's help and kindness when I need it most. I also want to express my heartfelt thanks to my former advisor Dr. Yangdo Choi for his longstanding supports and encouragement.

I had a good fortunate to spend time with many other friends and lab mate, Hong Luo during the courses of graduate school. In particular, I would like to thank AKPC people for their prayers and supports for me and my family. Although I served as a teacher, I also learned a lot from AKPC youth kids in my bible study classes, which make my life more plentiful and fruitful.

In this last moment of dissertation, I could not find exact words to express my gratefulness to my wife, Yumi. Since I met her at GB church, some day in October 2002, I have been in love with her. Her ceaseless supports, encouragement, prayer and love were the most important momentums in my journey. I thank you so much, and I love you. I am very thankful for my kids, Abigail Eunkyo and Christian Inkyo, who make Yumi and my life perfect. In addition, I would like to thank my parents, parents in law, brothers and sisters in Korea for their supports and trust.

Identification and Characterization of a Positive Control Region for Activation Induced Cytidine Deaminase Mediated Gene Conversion in Chicken B Cells

Publication No._____

Yonghwan Kim, Ph.D.

The University of Texas at Austin, 2009

Supervisor: Ming Tian

B cells have unique machinery to make up a large pool of antibody repertoire. After V(D)J recombination in early B cell development, the rearranged immunoglobulin genes are further diversified by somatic hypermutation (SHM), gene conversion (GC) and class switch recombination (CSR). Acitvation induced cytidine deaminase (AID) is a key initiating factor for SHM, GC and CSR. A majority of research data supports the model that AID modifies Ig genes at the DNA level by deaminating cytosines to uracils. The mutagenic activity of AID is largely restricted to Ig genes to avoid genomic instability in general. The specificity cannot be attributed to the primary sequence of the Ig genes since unrelated DNA is mutated by AID in the context of Ig genes. A clue to this problem is that AID function is dependent on transcription. Since not all transcribed genes are mutated by AID, there must be something special about the transcription of Ig genes, and the reasoning has prompted extensive analysis of Ig promoters and enhancers.

We addressed this question in chicken B cell line DT40. We identified a 2.4-kilobase regulatory region which is important for AID function both within and outside of Ig locus. This regulatory region contains binding sites for multiple transcription factors. Mutation of these binding sites impairs AID mediated gene conversion. In addition, ablation of NF-κB family member, c-Rel and p50, reduces the AID targeting function of this regulatory region. Since the implicated transcription factors have been reported to associate with histone acetylases, the regulatory region may function by facilitating the access of AID to target DNA. To test this hypothesis, we used the I-SceI endonuclease and dam methylase as probes for chromatin structure. We found that the regulatory region does not increase chromatin accessibility to these probes. In fact, the regulatory region appears to interfere with the cleavage of target DNA by I-SceI. Another possible role of the regulatory region could be direct recruitment of AID to Ig genes. To test this hypothesis, we utilized Dam identification method. Surprisingly, we found that the regulatory region facilitates AID targeting to the Igλ locus.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: Introduction

1.1. THE MECHANISMS OF ANTIBODY DIVERSIFICATION

B cells are capable of generating a diverse repertoire of antibodies to counter a wide variety of pathogens. The diversification of antibodies is accomplished through a programmed series of rearrangements and mutation events at the immunoglobulin (Ig) loci during B cell maturation. During the early stage of B cell development in the bone marrow, the Ig loci undergo V(D)J recombination to assemble functional genes that encode antigen receptors (Jung et al., 2006). Mature B cells emerge from the bone marrow, and circulate in the peripheral lymphoid tissues to monitor infectious agents (Janeway et al., 2005). B cells are activated upon antigenic challenge, and during B cell activation, the Ig loci undergo class switch recombination so that the B cell can produce different classes of antibodies with distinct effector functions. Contemporaneous with class switch recombination, the variable regions accumulate high levels of mutations (somatic hypermutation) that could improve the antigen binding affinity of the antibodies produced by the B cells. In avian species, the Ig variable region undergoes gene conversion as a means to diversify the primary repertoire of antigen binding sites during B cell maturation in the bursa (Ratcliffe, 2006; Reynaud et al., 1987). Following is a more detailed description of these four mechanisms of Ig diversification.

1.1.1. V(D)J recombination

An antibody molecule is composed of a pair of Ig heavy (IgH) and a pair of light (IgL) chains, and each Ig polypeptide can be further divided into variable and constant

regions. The variable regions of the IgH and IgL chains form the antigen binding site, while the constant regions of the IgH chain determine the effector functions of the antibodies and define antibody class or isotype (Jung et al, 2006; Janeway et al., 2005). One unique feature of Ig genes is that the functional coding sequences are assembled from variable (V), diversity (D) and joining (J) gene segments during B cell maturation, and this assembly process is accomplished through V(D)J recombination (Fig. 1.1). In the IgH loci, there are a large number of V, D and J gene segments. During the pro-B cell stage of B cell development, a D gene segment is first joined to a J gene segment, and subsequently the DJ gene segment is joined to a V gene segment in the IgH loci (Jung et al., 2006). The combined VDJ gene segments form the coding region for the variable region of IgH chain. V(D)J recombination is an ordered event in that D to J joining always takes place prior to V to DJ joining (Fig. 1.1). Moreover, the IgH locus undergoes V(D)J recombination prior to the IgL locus. After the completion of IgH V(D)J recombination, the B cell precursor progresses to the pre-B cell stage, and the IgL locus undergoes V(D)J recombination. The organization of the IgL loci is similar to the IgH locus in that there are a large number of gene segments involved in V(D)J recombination; the key difference is that there are no D gene segments at the IgL loci. The joining of V to J gene segments completes the assembly of the coding region for IgL. After successful assembly of the coding regions for IgH and IgL, B cell precursors express the antigen receptor on cell surface, and become immature B cells. After selection against self-reactivity, immature B cells migrate to the spleen and become mature B cells (Janeway et al., 2005).

V(D)J recombination is a site specific recombination process that is catalyzed by the Rag1 and Rag2 recombinases (Jung et al., 2006). The two proteins recognize a recombination signal sequences (RSS) at the border of the gene segments and introduce a

double stranded DNA break. Rag1 and Rag2 together form Rag endonuclease which induces DNA double strand breaks in the recombination signal sequence (RSS) located at the border of each segments. RSSs are composed of a highly conserved heptamer and nonamer which are separated by a spacer containing either 12 or 23 non-conserved nucleotides. V(D)J recombination only occurs between a 12-bp RSS and a 23-bp RSS, the restriction known as 12-23 rule (Jung et al., 2006). Repair of the double stranded DNA breaks by the non-homologous end joining system assembles two gene segments together (Rooney et al., 2004). The recombination process contributes to the diversity of antigen binding sites at several levels (Janeway et al., 2005). First, there are a large number of gene segments, and the random joining of these gene segments give rise to an enormous combinatorial diversity. Moreover, different IgH and IgL chains can be combined together to form distinct antigen binding sites, further increasing the range of potential antigen binding sites. Second, a unique feature of V(D)J recombination is that a hairpin is formed at the coding end of the gene segments after Rag mediated cleavage. The hairpin is opened with the endonuclease Artemis, which cleaves at random positions both within and around the hairpin to generate P nucleotides. Additionally, the ends are subject to the addition of random nucleotides by the terminal deoxynucloetide transferase (TdT) enzyme. These reactions lead to the formation of imprecise joints at the coding ends of the gene segments, and contribute substantially to the diversity of antigen binding sites.

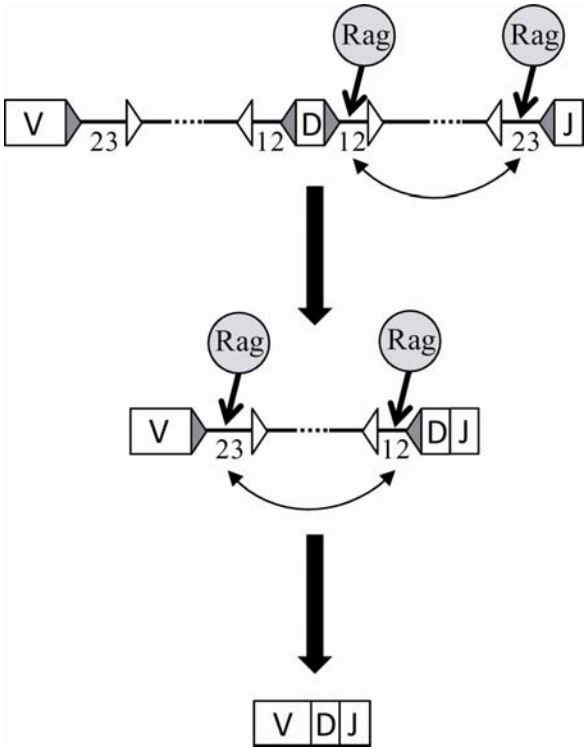

Figure 1.1. V(D)J recombination in IgH genes

V, D and J denote variable, diversity and joining segments, respectively. Recombination signal sequences (RSS) are shown at the border of each V, D and J segment. Gray triangles and blank triangles depict RS heptamers and nonamers, respectively. Spacer lengths are indicated below the line of RSS.

1.1.2. Gene conversion

V(D)J recombination is the principal mechanism to generate a diverse repertoire of antigen binding sites in mammals. By contrast, the Ig loci in avian species are composed of a limited number of functional gene segments, and V(D)J recombination alone is unable to create diverse antigen binding sites (Fig. 1.2*A*); instead, these species utilize Ig gene conversion to diversify their primary antibody repertoire (Reynaud et al., 1987, 1989). In the Ig loci of chicken, there are a large number of pseudo gene segments that precede the rearranged Ig variable regions. During B cell maturation, the Ig variable region undergoes gene conversion, which transfers random stretches of pseudo gene sequences into the rearranged Ig variable regions, and as a result, the range of antigen binding sites is substantially expanded. Another unique aspect of B cell maturation in avian species such as chicken is that B cell maturation takes place in the bursa, an organ with no mammalian counterpart (Arakawa and Buerstedde, 2004).

In contrast to V(D)J recombination, Ig gene conversion does not require the Rag1 and Rag2 proteins, but involve a distinct factor, called Activation Induced Cytidine Deaminase (AID) (Arakawa et al., 2002; Harris et al., 2002). The factor was first identified in mammalian B cells, and was later found to be essential for three distinct mechanisms of Ig diversification: class switch recombination, somatic hypermutation and Ig gene conversion. There are two views on how AID functions (Shivarov et al., 2009; Delker et al., 2009). In light of the homology of AID to the RNA editing enzyme, apolipoprotein B mRNA editing enzyme catalytic polypeptide 1 (APOBEC1), AID was initially proposed to function as an RNA editing enzyme that edits the mRNA of a factor involved in class switch recombination, somatic hypermutation and Ig gene conversion, but the putative editing substrate has not been identified so far (Muramatsu et al., 1999).

In the alternative model, AID was proposed to function as a DNA deaminating enzyme, converting cytidines in DNA into uridines (Di Noia and Neuberger, 2002, 2004). The subsequent processing of the uridines could lead to class switch recombination, somatic hypermutation, and gene conversion.

Avian B cells diversify their Ig repertoire by gene conversion and somatic hypermutation, but gene conversion is predominant. Inactivation of RAD51 paralogues (Sale et al., 2001; Bezzubova et al., 1997; Longerich et al., 2008; Yamamoto et al., 2005) or deletion of whole pseudo gene donors reduces gene conversion but induces somatic hypermutation (Arakawa et al., 2004). Although Ig gene conversion is obviously different from somatic hypermutation, both processes are likely initiated by AID mediated cytosine deamination in V region. The two processes diverge in subsequent processing of the primary lesion. Error-prone repair pathways are employed by somatic hypermutation to generate base substitutions, while homologous recombination is utilized to achieve gene conversion (Arakawa and Buerstedde, 2009). In the Igλ locus, only one set of V and J gene segment undergoes V(D)J recombination. The unique V region is diversified through gene conversion using upstream pseudo V gene fragments. In Ig gene conversion, AID deaminates cytosine in target DNA, leading to U:G mismatches. The uracil is excised by uracil DNA glycosylase (Di Noia and Neuberger, 2004; Di noia et al., 2007; Saribasak et al., 2006), and the resultant abasic site is further processed by either apurinic/apyrimidinic endonuclease (Arakawa and Buerstedde, 2009) or by the Mre11/Rad50/NBS1 complex to generate nicks in the Ig variable region (Yabuki et al., 2005; Larson et al., 2005). Homologous recombination machinery catalyzes the invasion of strand breaks into pseudo genes and copies pseudo gene sequences in the Ig variable region (Fig. 1.2*B*) (Sale et al., 2001; Arakawa et al., 2004).

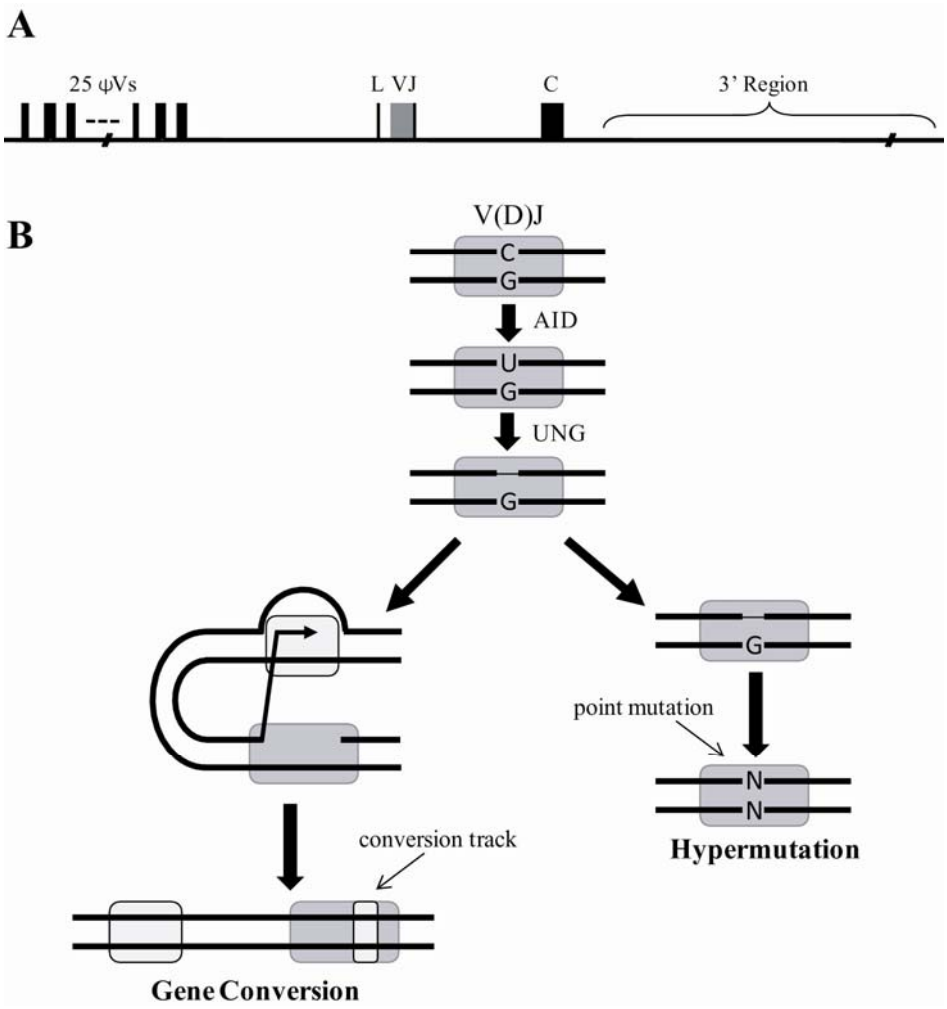

Figure 1.2. Ig diversification of in chicken B cells

(A) This Igλ map shows the positions of the relevant elements discussed in the text. L is the leader exon. VJ is the variable region consisting of joined V and J gene segments. C is the constant region. This diagram is drawn to scale. (B) Proposed model of Ig diversifications in avian B cells.

1.1.3. Somatic hypermutation

Encounter with antigens by B cells in the peripheral lymphoid tissues could lead to B cell activation. Some of the activated B cells undergo extensive proliferation and form a structure called germinal center in the secondary lymphoid tissues. Within the germinal center, the Ig variable regions of the activated B cells accumulate high levels of mutations, primarily point mutations and small deletions; the mutation rate of this process is estimated to be a million fold higher than the spontaneous mutation rate (Odegard and Schatz, 2006; Janeway et al., 2005). Some of these mutations could improve the antigen binding affinity of the antibodies, and result in the production of higher affinity antibody; for this reason, the process is called antibody affinity maturation. AID is essential for somatic hypermutation, and its mechanism of function can also be explained with the DNA deamination model (Fig. 1.3) (Muramatsu et al., 2000: Di Noia and Neuberger, 2007; Neuberger et al., 2005). The process is proposed to diverge from Ig gene conversion after the generation of uracils in the Ig variable region by AID (Neuberger et al., 2005). If DNA replication takes place over the uracil containing templates before its removal, the resultant daughter strands would contain C:G to T:A transition mutations. If DNA replication takes place over the abasic sites after the removal of uracil, normal replication machinery would be stalled at the abasic site, and translesion synthesis could take over, and insert random nucleotides in opposition to the abasic site (Fig. 1.3). As a result, the original C:G base pair could be mutated to different sequences. Additionally, the U:G mismatch could also be processed by the mismatch repair system (Rada et al., 1998; Phung et al., 1998; Frey et al., 1998; Wiesendanger et al., 2000; Bardwell et al., 2004), which could remove the mismatch in conjunction with a stretch of adjacent DNA. The gap could be filled in by DNA synthesis, which could also involve

error prone DNA polymerases to cause mutations (Zeng et al., 2001; Faili et al., 2004). In support of this model, inactivation of uracil DNA glycosylase leads to reductions in transition mutations (Di Noia and Neuberger, 2002; Rada et al., 2002; Imai et al., 2003), and deficiencies in certain error prone DNA polymerases have also been reported to affect the mutation spectrum of somatic hypermutation. Deficiencies in the mismatch repair system are associated with reduction of mutation in A:T base pairs. The exact reason for this phenotype is not fully understood; one potential explanation is that the error prone polymerase involved in the DNA synthesis step during mismatch repair tends to incorporate mutations at A:T base pairs. Overall, this experimental evidence is largely consistent with the key features of DNA deamination model for somatic hypermutation.

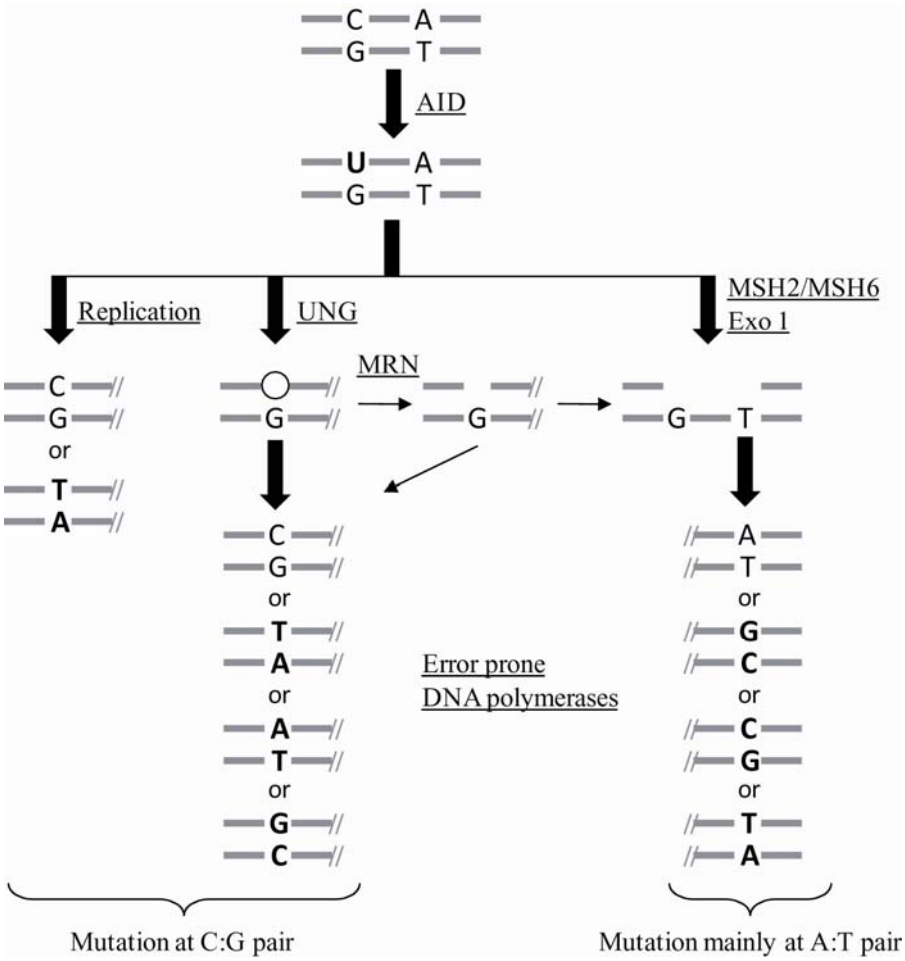

Figure 1.3. Illustration of a DNA deamination model of SHM

U:G mismatch triggered by AID undergoes different processes to generate mutations at C:G pairs and A:T pairs. Mutated nucleotides are indicated as bold characters. Possible factors involved in processes after AID are indicated as underlined.

1.1.4. Class switching recombination

Class switching recombination of immunoglobulin heavy chains also takes place during B cell activation (Janeway et al., 2005). In contrast to somatic hypermutation and gene conversion, class switch recombination involves the constant regions of IgH and alters the effector functions of antibodies (Chaudhuri and Alt, 2004). This reaction occurs between switch regions that precede each constant region except Cδ. These switch regions are composed of several kilobases of repetitive DNAs. It is thought that during B cell activation AID deaminates cytidines in the switch regions with high frequency. The subsequent removal of the uridines through uracil DNA glycosylase and apurinic/apyrimidinic endonuclease lead to strand breaks in the switch regions (Manis et al., 2002). Some of the strand breaks would occur on both strands and be adjacent to each other, resulting in the formation of double stranded DNA breaks, and ligation of the double stranded DNA breaks would give rise to the switch recombination products (Fig. 1.4) (Chaudhuri and Alt, 2004). Another pathway of processing uracils in the switch regions involves the mismatch repair system in analogy to the somatic hypermutation pathways, but it is unclear how the mismatch repair system would introduce strand breaks after the recognition of the U:G mismatch (Manis et al., 2002). In support of this model, deficiencies in the uracil DNA glycosylase result in severe reductions in class switch recombination (Rada et al., 2002), and ablation of components of the mismatch repair systems also impair class switch recombination (Schrader et al., 1999). The combined ablation of both uracil DNA glycosylase and essential factors for mismatch repair pathways abolishes class switch recombination, which is entirely consistent with the predictions of the DNA deamination model of class switch recombination.

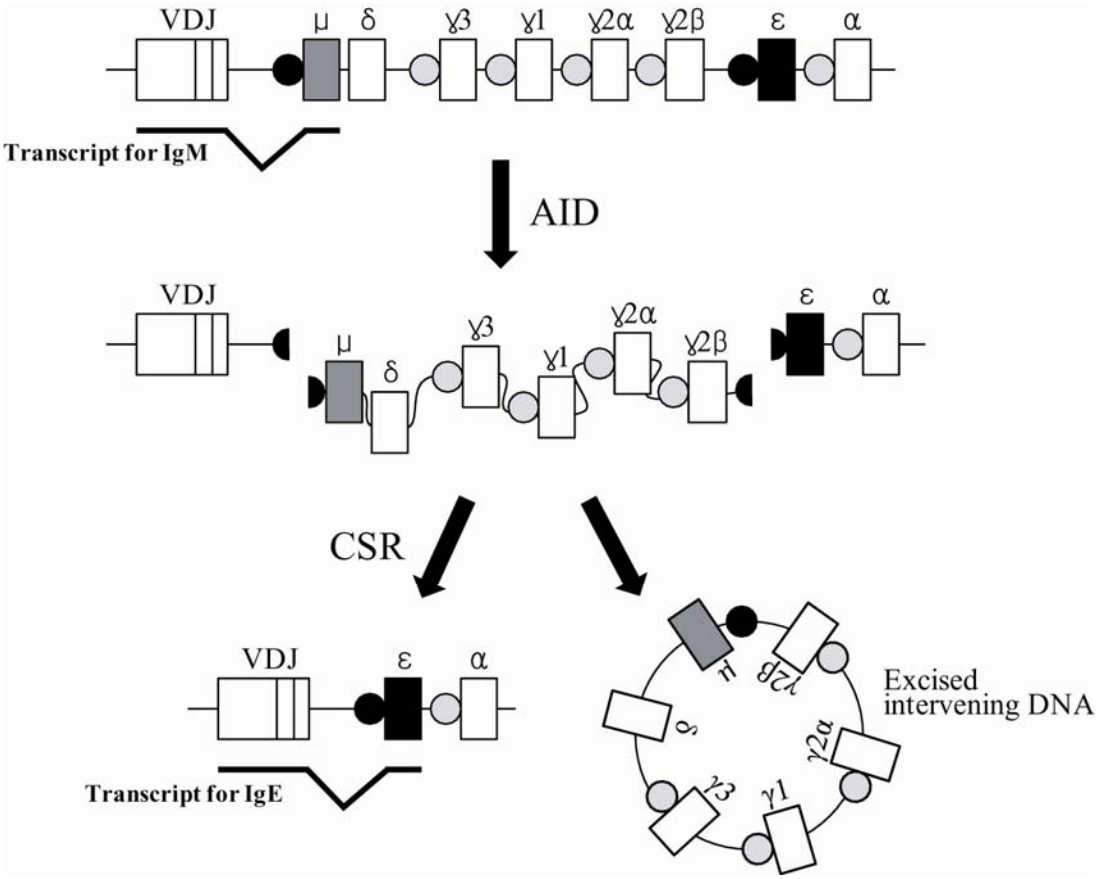

Figure 1.4. Illustration of class switch recombination at the mice IgH locus

Mice IgH constant region is depicted with rearranged variable region. In this diagram, B cell is stimulated to undergo class switching to Cε.

1.2. REGULATION OF ACTIVATION INDUCED CYTIDINE DEAMINASE

Although AID is essential for three important mechanisms of antibody diversification, it also poses a potential danger to the host: the mutagenic activity of AID could potentially cause deleterious mutations in the genome (Odegard and Schatz, 2006). Therefore, the question arises as to how the mutation activity of AID is controlled to prevent genomic instability. Given the importance of this regulation, AID function is controlled at multiple levels. First, AID expression is restricted to B cells undergoing Ig diversification: activated B cells (Muramatsu et al., 1999) and bursa B cells (Arakawa and Buerstedde, 2009). It has been shown that ubiquitous expression of AID in transgenic mice leads to mutations in other cell types and induces tumor development (Okazaki et al., 2003; Takizawa et al., 2008; Robbiani et al., 2008)). The second level of control is that AID is distributed in both the cytoplasm and the nucleus (Barreto et al., 2005). The N-terminus of AID contains a nuclear localization signal, while the C-terminus of the protein contains a nuclear export signal (Ito et al., 2004). At steady state, most of the AID protein is located in the cytoplasm to prevent potential mutagenic activity on nuclear DNA. It has been shown that removal of the C-terminal export signal leads to elevated AID levels in the nucleus and high mutation rates (Shinkura et al., 2004; Barreto et al., 2003; Ta et al., 2003). The third level of control is the confinement of AID activity to the Ig loci. Even with transcriptional control and differential sub-cellular location, AID needs to be present at sufficient levels in B cells to mediate Ig diversification, and its activity needs to be confined to the Ig targets to prevent mutations in other parts of the genome. Although certain non-Ig genes are also subject to some level of AID mediated mutation events (Shen et al., 1998; Pascualucci et al., 1998, 2001; Gordon et al., 2003; Muschen et al., 2000; Liu et al., 2008), the mutation

frequencies in these non-Ig targets are substantially lower relative to those sustained by the Ig genes. Currently, it is still largely unclear how AID activity is targeted to the Ig loci. The fourth level of control of AID function is that there appears to be differential repair of AID mediated lesions in Ig versus non-Ig loci. It has recently been reported that AID in fact can introduce mutations in a large number of genomic locations outside of the Ig loci, but most of the untargeted mutations are removed by the high fidelity repair systems (Liu et al., 2008; Liu and Schatz, 2009). The evidence for the conclusion is that many genes accumulate mutations when uracil DNA glycosylase and mismatch repair systems are disabled (Liu et al., 2008). It has been proposed that the repair process at the Ig loci might have been corrupted in some way to prevent removal of AID mediated lesion by the high fidelity repair systems, but the exact mechanism behind this proposal is presently unclear.

1.3. TRANSCRIPTION IS ESSENTIAL FOR AID FUNCTION IN SHM, GC AND CSR

Transcription is a key requirement for all aspects of AID mediated recombination and mutation events (Fukita et al., 1998; Bachl et al., 2001). The requirement may reflect the substrate requirement of AID: the enzyme can only deaminate cytidines in the context of single stranded DNA. During transcription, the DNA template is transiently unwound, and this provides a window of opportunity for AID function (Peters and Storb, 1996; Rada et al., 1997; Shen et al., 2009). However, certainly not all the transcribed genes are subject to AID mediated deamination, so the question remains as to why transcription at the Ig loci allows AID function. It has been hypothesized that the transcription at the Ig loci may be special in certain way: for example, the transcription factors involved in Ig gene transcription may have the ability to recruit AID to facilitate

their interaction with the Ig targets that are undergoing transcription. For this reason, there has been extensive analysis of Ig transcription elements and factors in the hope of identifying the key elements and factors in this process. Most of these studies were carried out in mice. The consensus of these studies is that the Ig promoter is not the determinant for AID functional specificity since non-Ig promoters can replace Ig promoters and induce somatic hypermutation (Betz et al., 1994; Tumas-Brundage and Manser, 1997). The analysis of Ig enhancers is complicated by the fact that there are multiple enhancer elements with partially overlapping functions at the murine Ig loci, and deletions of individual enhancer elements do not have a major impact on AID function at the endogenous loci (van der Stoep et al., 1998; Morvan et al., 1994; Inlay et al., 2006; Perlot et al., 2005; Xiang and Garrard, 2008; Ronai et al., 2005). Similarly, analysis of the chicken Ig loci that undergo AID mediated gene conversion also implicates the presence of complex regulatory elements involved Ig gene conversion (Kothapalli et al., 2008). To address this question, we carried out a systematic analysis of the regulatory elements at the chicken Igl locus to characterize the elements that could target AID function.

1.4. *CIS*-ACTING AID TARGETING ELEMENTS

There is tight linkage between transcription and the AID function: transcription is required for the AID activity and transcription rate is largely correlated to mutation frequency. However, the primary Ig variable region and promoters do not attribute to the specific targeting of AID to Ig loci in general. These findings lead to the hypothesis of which Ig enhancers play a role in targeting AID to Ig loci (O'Brien et al., 1987). In an effort to identify the Ig enhancers, the Igκ transgene, which is driven from the

rearranged Igκ light chain, has been used in early studies. Since the transgene undergoes SHM upon B cell activation, it became a good model system to identify the Ig enhancer function on SHM. Two enhancer elements have been found in the mice Igκ light chain; one is the iEκ enhancer located in JC intronic region and the other is the 3'Eκ enhancer (Betz et al., 1994). Without 3'Eκ, both transcription rate and SHM frequency were reduced. Since transcription rate is correlated to SHM, the role of 3'Eκ on targeting AID is not clear. Interestingly, deletion of iEκ in the transgene context reduced SHM activity but maintained transcription, suggesting that iEκ play a role in targeting AID activity. However, the same is not true in the endogenous Igκ locus (van der Stoep et al., 1998; Gorman et al., 1996; Inlay et al., 2006): ablation of either iEκ or 3'Eκ was dispensable for transcription and SHM. These discrepancies of the enhancer activities between Igκ transgene and endogenous Igκ locus indicate that there could be unknown redundant elements in the endogenous Igκ locus, which makes the analysis of enhancer more complicated.

In the same way, an enhancer in chicken Igλ locus, initially identified in transgenic and transfection experiments (Lauster et al., 1993; Bulfone-Paus et al., 1995), was later found to be largely dispensable for both gene conversion and transcription in the endogenous Igλ locus (Yang et al., 2006; Kothapalli et al., 2008), and the region downstream of the enhancer was reported to be important, instead. Another study identified multiple regulatory elements in the Igλ locus (Blagodatski et al., 2009). Notably, most of the immunoglobulin enhancers contain potential *trans*-acting factor binding sites for IRF4, E2A proteins, NF-κB and Mef2 (Schoetz et al., 2006; Combriato and Klobeck, 2001; Satyaraj and Storb, 1998; Klein et al., 2006). In fact, it was reported that E2A expression stimulates SHM and GCV (Schoetz et al., 2006). Especially, in case of chicken, Igλ light chain 3' core enhancer contains potential binding

site for all these four *trans*-acting factors (Bulfone-Paus et al., 1995). These factors are expressed in germinal center B cells. These findings suggest that there might be interactions between the *trans*-acting factors and *cis*-acting elements in the enhancer, resulting in targeting of SHM and GCV to the Ig locus.

CHAPTER 2: Materials and Methods

2.1. GENE CONVERSION SUBSTRATES

The gene conversion reporter was generated by inserting an I-SceI recognition sequence into the SalI site of the puromycin resistance gene (*puro*): GTC Ggc tag gga taa cag ggt aat gcA CCC (the inserted sequences are in lower case letters). A fragment of *puro* serves as the donor in gene conversion; if the A of the start codon is designated as position +1, the fragment corresponds to +1 to +333 of the open reading frame. The *puro* fragment is positioned 2-kb upstream of the promoter.

Three promoters were used to transcribe the *puro* reporter. If the transcription initiation site is designated as +1, the Igλ promoter contains sequences from -1,363 to +25. The chicken EF-1α promoter contains sequences from -315 to +25. A fragment of the first intron (+139 to +779) of EF-1α gene was cloned upstream of the promoter since this region shows enhancer activity. The chicken β-actin promoter contains sequences from -276 to +10, and a fragment of the first intron (+11 to +959) was cloned upstream of the promoter to stimulate transcription.

The following description refers to the rearranged Igλ locus of DT40 and the transcription initiation site of the Igλ promoter is designated as +1. The gene conversion substrate was integrated at position +4,007. Region A and B encompass sequences from +4,007 to +6,449 and +6,450 to +14,124, respectively. Element A1 corresponds to position +5280 to +5721. Element A2 is from +5722 to +6449. Given the large number of targeting constructs involved in this study, details of the constructs and recombination strategy are omitted.

2.2. p50 AND c-REL KNOCKOUT

The transcription initiation site of p50 gene is defined as position +1 (G^{+1}AGGTCGCGC). The first targeting construct eliminates +30,071 to +34,252, while the second targeting construct deletes +30,071 to +32,438. The transcription initiation site of c-Rel is defined as position +1 (G^{+1}CGAGGTGCG). The first targeting construct eliminates +2,996 to +12,258, while the second targeting construct deletes +4,495 to +9,346. Details of the targeting construct are omitted.

2.3. TRANSFECTION

All the cell lines in this study were derived from DT40cre1 (Arakawa et al., 2001). The cells were cultured in RPMI media supplemented with 10% fetal bovine serum, 1% chicken serum, 100 units/mL penicillin, 100 mg/mL streptomycin and 28 μM β-mercaptoethanol. For transfections, 10 μg of linearized targeting construct was electroporated into $10x10^6$ cells. Drug selection started a day after transfection. For puromycin selection, the drug was used at a concentration of 0.5 μg/mL. For *gpt* selection, mycophenolate was added to a final concentration of 5 μg/mL. Stable clones were analyzed with Southern blotting to detect homologous recombination events. Subsequently, the drug selection marker was removed through flanking loxP sites, and the recombination is catalyzed by cre recombinase expressed from a tamoxifen inducible transgene in DT40cre1 cell line (Arakawa et al., 2001). After one day of tamoxifen treatment, the cells were subcloned, and deletion of the drug selection marker was confirmed with Southern analysis.

2.4. GENE CONVERSION ASSAY

The cells were subcloned by serial dilution in 96-well tissue culture plates. The day of subcloning was day 0. After 9 days, single colonies were transferred into 24-well tissue culture plate and cultured for 3 more days. At day 12, a portion of the cells was transferred into 12 mL media containing 0.5 μg/mL puromycin, while another aliquot of the cells was diluted into 12 ml media without puromycin. The cell numbers in each part were indicated in the Table 1 through 7. These numbers were chosen because they yield countable numbers of colonies in most of the experiments. Each 12 mL of cells was distributed into a 96-well tissue culture plate. Colonies were counted under microscope 6-7 days later.

2.5. I-SCEI INDUCIBLE CELL LINE

I-SceI is expressed from a tetracycline-inducible system. The cDNA for the reverse tetracycline repressor (rTetR) was under the control of the chicken β-actin promoter, and the expression construct was stably integrated into $AID^{-/-}$ DT40 cell line (Arakawa et al., 2002). The cDNA for I-SceI was cloned downstream of the Tet-inducible promoter, and the construct was stably integrated into the cell line that expresses rTetR.

2.6. DAM ACCESSIBILITY ASSAY

The cDNA of E.coli DNA methyltransferase (Dam) was cloned by polymerase chain reaction. A tamoxifen inducible Dam expression construct was designed. In this construct, the *dam* gene is cloned downstream of the coding region for gpt selection marker, which is flanked with loxP sites. Deletion of the selection marker via cre-

mediated recombination, induced by treating DT40 cells with tamoxifen, removes the selection marker and places the dam gene under the control of the chicken β-actin promoter. We integrated the inducible dam construct into the AID locus. Genomic DNAs were harvested after 0, 8, and 12 hours of Dam induction and digested by methylation sensitive endonuclease MboI with a combination of other endonucleases. The extent of DNA methylation was analyzed by Southern blotting and quantified by Phosphoimager.

2.7. DAM IDENTIFICATION (DAMID)

E.coli dam gene was fused to AID cDNA isolated from DT40 cells. A tamoxifen inducible AID-Dam expression construct was developed, as described above. We integrated the inducible AID-Dam construct into the AID locus of both DT40 clones, L1 and L2. Genomic DNAs were harvested after 0, 8, and 12 hours of induction and digested by EcoRI, XbaI and MboI for analyzing the DNA methylation of Igλ locus and BamHI, BglII and MboI for EF-1α locus. The extent of DNA methylation was analyzed by Southern blotting and quantified by Phosphoimager.

2.8. REVERSE TRANSCRIPTASE (RT)-PCR

Grow DT40 cells to confluent in an either 24 or 12 well plate. Isolate total RNA. Before setting up RT reaction, denature 3 μL RNA with 50 pmol oligo dT for 5 min at 65°C, chill it on ice for 2 min and spin it down. Reaction for the first strand includes 1X RT buffer, 0.01 M DTT, 0.5 mM dNTP and reverse transcriptase in 20 μL. Keep the reaction for 1 hour at 37°C. After reaction, take 2 μL (out of 20 μL) and add 18 μL of TE (one to ten dilutions). Use 1 μL of diluted first strand for PCR reaction.

2.9. SOUTHERN BLOT ANALYSIS

Grow DT40 cells to confluence in a 24 well plate. Transfer cells into 1.5 mL tube. Spin down cells at 3000 rpm for 5 min, remove supernatant, and resuspend cells with 250 μL PBS. Add 250 μL cell lysis buffer with 1 μL of 10 mg/mL proteinase K. Keep the tube at 65°C incubator overnight. Next day, add same volume of propanol, and shake the tube (you will see DNA thread). Spin it down at Max for 5 min. Remove supernatant and wash genomic DNA with 70% ethanol. Resuspend genomic DNA in 100 μL TE at 65°C incubator overnight. Take 50 μL of the resuspend DNA and set up restriction endonuclease reaction in 200 μL. Digest genomic DNA overnight at 37°C. Next day, add 500 uL 100% ethanol and spin down the digested DNA. Resuspend DNA in 30 uL TES by keep the DNA in 65°C incubator for about 5 hours. Size-fractionate the completely resuspended DNA in 0.8 ~ 1.2% agarose gel overnight. Next day, denature the gel in 0.5 N NaOH for 1 hour, and transfer the DNA onto nylon membrane overnight. Next day, fix the blot by baking it at 80°C vacuum incubator for 1 hour. Hibridize the blot with appropriate probe.

2.10. NORTHERN BLOT ANALYSIS

Grow DT40 cells to confluence in an either 24 or 12 well plate. Transfer cells into 1.5 mL tube. Spin down cells at 3000 rpm for 5 min, remove supernatant, and resuspend cells in 500 μL RNA extraction solution* (see below). Add 100 μL of chloroform and voltex it. Spin down for 15 min at 12 rcf. Take supernatant (about 300 μL) and add 70 % propanol. Spin down for 15 min at 12 rcf and wash the RNA with 70% ethanol. Resuspend RNA in 10 ~ 15 uL ddH_2O. Size-fractionate the

isolated RNA in 0.8% denaturing agarose gel. Next day, transfer RNA onto nylon membrane overnight. Next day, fix the RNA on nylon membrane by baking the blot at 80°C vacuum incubator for 1 hour. Hybridize the blot with appropriate probes.

* Recipe for 50 μL of RNA extraction solution
  Add 20 mL of 4M Guanidine-CNS with 0.5% of Sarcosyl
  Add 2 mL of 2M NaOAc pH 4
  Add 140 μL of β-mercaptoethanol
  Mix it gently by invert down
  Add final 20 mL of H_2O saturated Phenol pH 4
  Mix it gently by invert down

CHAPTER 3: Identification of an Important Regulatory Region for AID Mediated Gene Conversion in Chicken B Cell Line DT40 [1]

3.1. INTRODUCTION

Activation induced cytidine deaminase (AID) is a key initiating factor for somatic hypermutation (SHM) and gene conversion (GC), and the related process, class switching recombination (CSR) (Arakawa et al., 2002; Muramatsu et al., 2000; Revy et al., 2000). The majority of research evidence supports the model that AID catalyzes deamination of cytosines to uracils in target DNA (Di Noia and Neuberger, 2002; Di Noia and Neuberger, 2007). In the case of gene conversion, this dG/dU mismatch is processed by uracil DNA glycosylase and apurinin/apyrimidinic endonuclease, generating a nick in the Ig variable region (Di Noia and Neuberger, 2002; Sale, 2004). Homologous recombination machinery induces the invasion of the strand break into pseudo genes and copies pseudo gene sequences into the Ig variable region (Sale et al., 2001; Arakawa et al., 2004). Because of mutagenic activity, confinement of AID to Ig loci is critical for maintaining genome integrity (Odegard and Schatz, 2006; Lebecque and Gearhart, 1990).

Transcription is required for AID activity, and the rate of SHM and GC is largely correlated to the level of transcription (Fukita et al., 1998; Bachl et al., 2001). In addition, somatic hypermutation is restricted within 1 1.5-kb region from the transcription start site of Ig genes (Lebecque and Gearhart, 1990). It is generally believed that neither the promoter nor the rearranged Ig variable primary sequences are critical for AID targeting (Betz et al., 1994; Yelamos et al., 1995). These findings and

[1] *Reproduced in part from Kim and Tian (2009) Mol Immunol. (46) 3283-3291. Appropriate permission has been granted by the publisher.*

the tight linkage between AID function and transcription have prompted intense research of Ig gene enhancer for their roles in targeting AID to Ig loci (Odegard and Schatz, 2006). Deletion of intronic enhancer (iEκ) in the Igk transgene in mice showed severe reduction of SHM while maintaining transcription level as wild type (O'Brien et al., 1987; Betz et al., 1994). However, removal of the iEκ in the endogenous context demonstrated that iEκ is not essential for both SHM and transcription (van der Stoep et al., 1998; Gorman et al., 1996; Inlay et al., 2006). This discrepancy can be explained by redundant elements in the endogenous Ig locus because the transgene only contains minimal components required for inducing SHM. To date, little is known about the molecular mechanism of targeting AID. Obviously, additional study is required to characterize the elements that target AID to Ig loci. In chapter I, I analyzed the regulatory region which is important for AID mediated GC.

3.2. RESULTS

3.2.1. Establishment of a novel gene conversion assay system

Two GC assay protocols have been widely used (Arakawa et al., 2002, 2004; Harris et al., 2002; Sale et al., 2001). These two methods are based on the same concept: Ig gene conversion frequency is reflected by the changes of surface IgM expression, detected by flow cytometry. However, the way of measuring the GC frequency is opposite (described below). One GC assay system uses wild type DT40 cells carrying surface IgM. Since some portions of GC events can cause deleterious mutations in variable regions, the GC frequency can be analyzed by looking at the resultant DT40 cells losing surface IgM. The other GC assay system uses a DT40 cell

line which does not express surface IgM due to the frameshift mutation on its V region. This mutated V region can revert to wild type by certain GC events to generate surface IgM positive DT40 cells. However, these two GC assays take a long time, more than 3 weeks of cell culture, to have enough surface IgM changes detectable by the flow cytometry. This long term culture could reduce the credibility of these mutation assays because early GC events might produce more surface IgM alterations than late GC events.

To measure GC frequency in a more sensitive way, we have developed a novel GC assay system in DT40 cells (Fig. 3.1). We applied a positive drug selection scheme to our assay. The reporter was made of puromycin resistant gene (*puro*) which is inactivated by inserting 18-bp Yeast I-SceI recognition sequence. The mutation can be eliminated by gene conversion using wild type *puro* fragment located in 2-kb upstream of promoter region (Fig. 3.1). Therefore, DT40 clones, which undergo GC, form colonies in puromycin-treated culture. Point mutations, typical of somatic hypermutation, are unable to restore gene function. As a result, puromycin resistance selects specifically for gene conversion events. Our novel GC assay system is less time consuming, is more sensitive than any other GC assay method developed, and is stable over the many generations in DT40 cells. More importantly, since the readout of the assay is not dependent on Ig expression, the substrate enables us to study gene conversion outside of the Ig loci. In addition, transcriptional elements can be easily manipulated in the GC substrate.

The substrate was integrated into the Igλ locus via homologous recombination (Fig. 3.2*A*). The *puro* reporter is under the control of the Igλ promoter (L1, Fig. 3.2*B*), and robust transcription of the reporter was observed, presumably due to the influence of strong Ig enhancers (L1, Fig. 3.3*A*). To analyze gene conversion, the cell line was

subcloned by dilution to eliminate puromycin resistant cells that were generated by ongoing gene conversion in culture (Buerstedde et al., 1990; Kim et al., 1990). After 12 days, progenies of a single cell were distributed into 96-well tissue culture plates in media with puromycin. As a control for plating efficiency, a fraction of the cells were cultured in media without puromycin. The number of colonies was counted under microscope 6-7 days later. For purpose of illustration, we define the number of cells distributed into the puromycin media and control media as P and C, and the resulting colony numbers as P' and C', respectively. The gene conversion frequency was calculated by following equation.

$$GC\ frequency = \frac{P'/P}{C'/C}$$

Applying this method to L1, the median gene conversion frequency from 6 experiments was 3.4×10^{-3} (L1, Fig. 3.3*B* and Table 1). Based on Southern analysis, all the puromycin resistant clones have regained the SalI site destroyed by the mutation in the *puro* reporter, and represent gene conversion products (Fig. 3.4). To ensure that gene conversion was mediated by AID, the substrate was introduced into AID deficient cells (L8, Fig. 3.2*B*). As shown in Figure 3.3 and Table 1, no puromycin resistant colonies were generated in the absence of AID even though the *puro* reporter was transcribed normally (L8, Fig. 3.3*A*). Thus, the substrate serves as a valid assay for AID-mediated gene conversion.

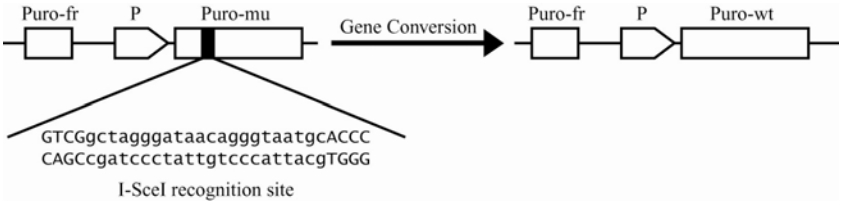

Figure 3.1. Gene conversion assay

This diagram illustrates the gene conversion substrate. Puro-fr is the *puro* fragment that serves as the donor in gene conversion. P represents promoter. Puro-mu is the mutated *puro* reporter; the mutation is represented with a black bar and inserted nucleotides are shown below. Puro-wt is the wild-type *puro* that is generated through gene conversion.

3.2.2. Identification of a regulatory region for AID mediated gene conversion

To locate critical regions for AID-mediated gene conversion, we performed a sequential deletion analysis of the Ig 3' enhancer. As shown in Figure 3.2*A*, we deleted the 2.4-kb (L2, Fig. 3.2*B*) and 7.4-kb (L3, Fig. 3.2*B*) enhancer regions adjacent to the GC substrate by homologous recombination. The 2.4-kb and 7.4-kb region are designated as Region A and Region B, respectively (Fig. 3.2*A*). Region A and Region B deletion reduced GC frequency by 15 fold and 6 fold, respectively (compare L1, L2 and L3, Fig. 3.3*B* and Table 1). Since GC frequency is correlated to transcription level, we performed Northern blot analysis of the GC substrate to measure transcripts (Fig. 3.3*A*). Both Region A and Region B deletion showed about 70% of transcription level compared to wild type. Therefore, the reduction of GC activity cannot be explained by transcriptional impairment, suggesting that the Igλ 3' region plays a role in targeting AID-mediated gene conversion. GC assay results show that the *cis*-acting elements involved in AID recruiting is redundantly dispersed throughout the 3' region, but Region A, which is closer to the GC substrate, is more critical for AID mediated GC in the endogenous Ig locus.

Since deletion A has a more pronounced phenotype, we extended the analysis of this region to substrates under the control of chicken EF-1α and β-actin promoters (L4 through L7, Fig. 3.2*B*). Both promoters function comparably to Igλ promoter with respect to gene conversion and transcription, and deletion A resulted in a similar phenotype (Fig. 3.3*A* and *B* and Table 1). Clearly, Region A is important for gene conversion, and its function is beyond transcriptional activation. It was reported that the human EF-1α promoter functions poorly for Ig diversification at the Igλ locus of DT40

cells (Yang et al., 2006). In our study, the chicken EF-1α promoter was used, and the difference may underlie the discrepancy.

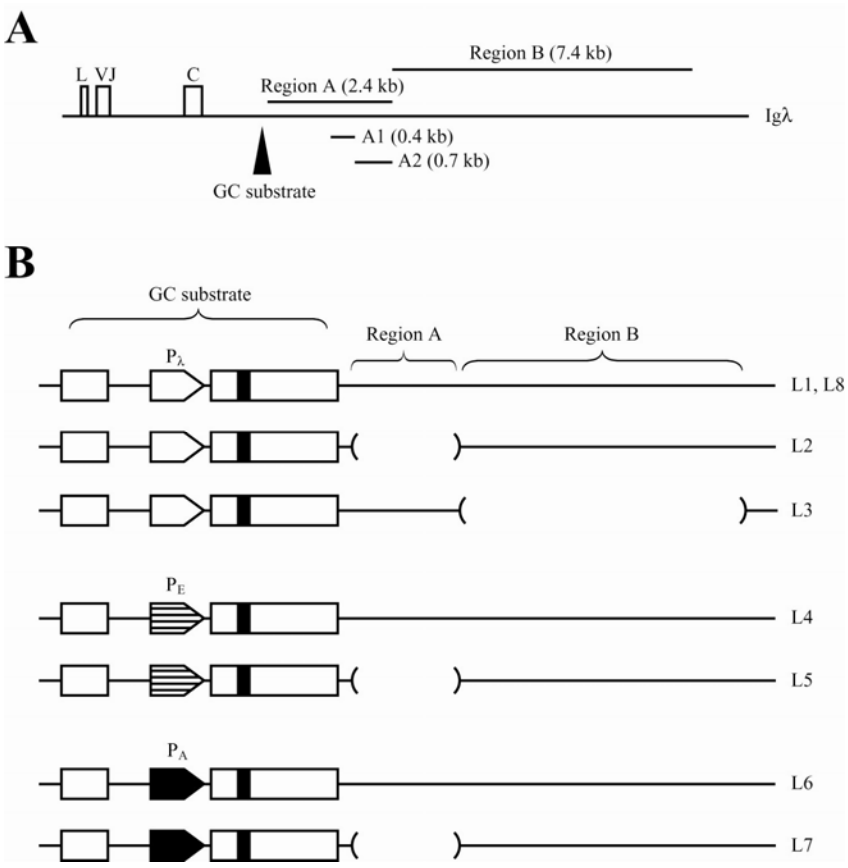

Figure 3.2. Gene conversion substrates in the endogenous Igλ locus

(A) This Igλ map shows the positions of the relevant elements discussed in the text. L is the leader exon. VJ is the variable region consisting of joined V and J gene segments. C is the constant region. The integration site of gene conversion (GC) substrate is indicated with an arrowhead. This diagram is drawn to scale. (B) This diagram illustrates the GC substrates (L1-L8) in the Igλ locus. P_λ, P_E, and P_A represent the Igλ promoter, EF-1α promoter, and β-actin promoter, respectively. Deletion of Region A or B is indicated with a bracket. L8 is an $AID^{-/-}$ DT40 cell line; all the other substrates are in $AID^{+/+}$ DT40 cells. This diagram is not drawn to scale.

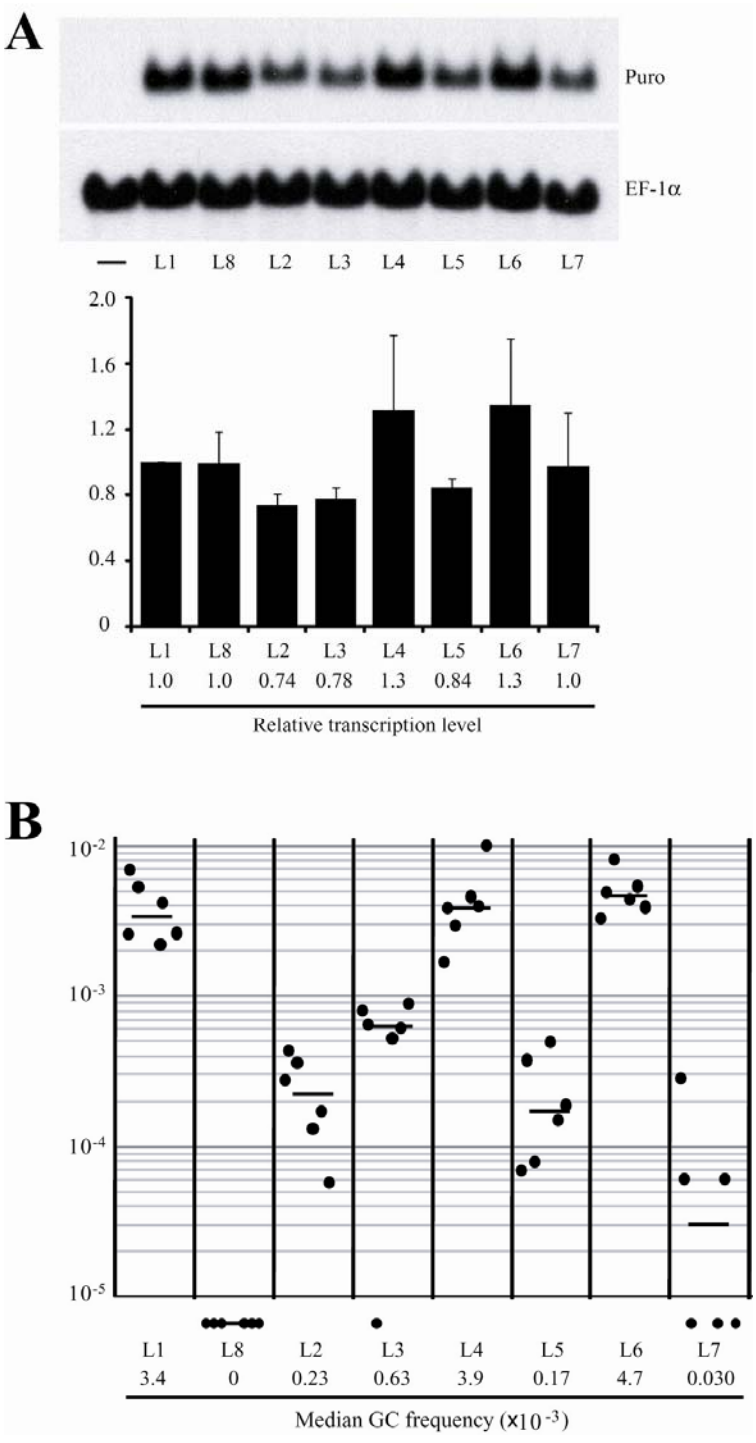
A
Puro
EF-1α
— L1 L8 L2 L3 L4 L5 L6 L7
2.0
1.6
1.2
0.8
0.4
0
L1 L8 L2 L3 L4 L5 L6 L7
1.0 1.0 0.74 0.78 1.3 0.84 1.3 1.0
Relative transcription level
B
10-2
10-3
10-4
10-5
L1 L8 L2 L3 L4 L5 L6 L7
3.4 0 0.23 0.63 3.9 0.17 4.7 0.030
Median GC frequency (×10-3)

Figure 3.3. Gene conversion analysis at the Igλ locus

(A) This figure shows the transcription levels of gene conversion substrates. The upper autoradiograph is a Northern blot hybridized with a *puro* probe. The lower autoradiograph is the same Northern blot hybridized with an EF-1α cDNA probe. The source of RNA is indicated below the autoradiograph. The first lane (-) contains RNA from DT40 cells without gene conversion substrate. The histogram compares the relative levels of *puro* transcript expressed from L1 through L8. Transcript level is based on quantification of hybridization signals of the Northern blot with PhosphoImager. A ratio of *puro*/EF-1α was calculated for each sample, and the value for L1 is set at 1.0. The histogram represents the average of three experiments, and the average value is written below the graph. Error bars represent standard deviations. (B) The GC frequencies of L1 through L8 are shown in this plot. Each dot represents the data from one experiment. The median value is indicated with a bar and written below the plot.

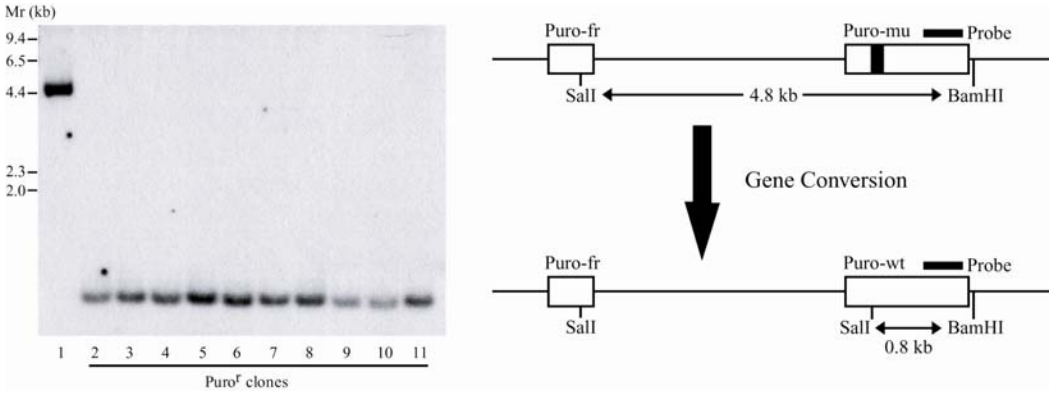

Figure 3.4. Regeneration of SalI site by gene conversion

The autoradiograph is Southern analysis of DNAs isolated from total L1 cells (lane 1) and puromycin resistant (puro^r) colonies (lanes 2-11). The DNAs were digested with SalI and BamHI. The DNA size markers (M) are indicated to the left of the autoradiograph. The probe and the expected restriction fragments are illustrated in the diagram to the right of the autoradiograph.

Substrate	Puro[a]	Control[b]	GC[c]	Substrate	Puro[a]	Control[b]	GC[c]
L1	42	81	2.6×10^{-3}	L5	1	76	7.0×10^{-5}
	106	76	7.0×10^{-3}		5	66	3.8×10^{-4}
	60	56	5.4×10^{-3}		1	66	8.0×10^{-5}
	31	70	2.2×10^{-3}		6	60	5.0×10^{-4}
	62	74	4.2×10^{-3}		2	66	1.5×10^{-4}
	42	80	2.6×10^{-3}		3	79	1.9×10^{-4}
L2	4	72	2.8×10^{-4}	L6	52	78	3.3×10^{-3}
	5	58	4.3×10^{-4}		68	69	4.9×10^{-3}
	5	70	3.6×10^{-4}		104	63	8.3×10^{-3}
	2	76	1.3×10^{-4}		64	72	4.4×10^{-3}
	2	58	1.7×10^{-4}		95	87	5.5×10^{-3}
	1	86	5.8×10^{-5}		64	81	4.0×10^{-3}
L3	12	74	8.1×10^{-4}	L7	4	71	2.8×10^{-4}
	7	54	6.5×10^{-4}		1	81	6.0×10^{-5}
	0	61	0		0	61	0
	7	67	5.2×10^{-4}		0	74	0
	8	65	6.2×10^{-4}		0	77	0
	10	56	8.9×10^{-4}		1	86	6.0×10^{-5}
L4	25	74	1.7×10^{-3}	L8	0	54	0
	54	70	3.9×10^{-3}		0	74	0
	34	58	2.9×10^{-3}		0	42	0
	67	73	4.6×10^{-3}		0	73	0
	49	62	4.0×10^{-3}		0	85	0
	174	65	1.3×10^{-2}		0	82	0

Table 1. Gene conversion at Igλ locus.

The data in each row represent analysis of one colony.
a. Puromycin resistant colony per $2x10^4$ cells
b. Colony number per 100 cells in media without puromycin
c. Gene conversion frequency = $(a/2x10^4)/(b/100)$

3.2.3. Region A is sufficient for targeting gene conversion outside of Ig locus

A more stringent test of Region A function is to see whether it is sufficient to target AID activity outside of Ig locus. In this experiment, we put the *puro* reporter under the control of Igλ promoter plus Region A, while EF-1α promoter serves as a control (Fig. 3.5*A*). These substrates were randomly integrated into the genome of DT40 cells, and a single copy integrant from each substrate was chosen for analysis. Since the integration sites are unknown, the substrates are labeled as X and Y, and the promoters are indicated with the subscript λ and E, respectively (Fig. 3.5*A* and *B*). Under the control of Igλ promoter plus Region A, substrate X_λ underwent gene conversion at a median frequency of 3.7×10^{-4} (X_λ, Fig. 3.6*B* and Table 2). By comparison, 41-fold lower gene conversion activity was observed when the reporter was transcribed with EF-1α promoter (Y_E, Fig. 3.6 and Table 2). Since these two substrates likely reside in different genomic locations, the result could reflect position effects. To address this issue, we replaced the Igλ promoter and Region A in substrate X_λ with EF-1α promoter through homologous recombination (X_E, Fig. 3.5*A*). Conversely, the EF-1α promoter in Y_E was changed into Igλ promoter plus Region A (Y_λ, Fig. 3.5*B*). The swapping experiment allows comparison of the two promoters at the same genomic location. As shown in Fig. 3.6*B* and Table 2, the introduction of Igλ promoter and Region A into Y locus (Y_λ) elevated gene conversion frequency by 6-fold. Reciprocally, gene conversion activity at the location X was essentially lost when the reporter was transcribed with EF-1α promoter (X_E, Fig. 3.6*B* and Table 2). Thus, Igλ promoter plus Region A is capable of recruiting AID to at least two genomic locations that are not AID targets under the control of EF-1α promoter. Relative to Igλ promoter and Region A, EF-1α promoter appears weaker (50% at position X and 70% at position Y) and produces a different transcript (puro*, Fig. 3.6*A*), which may arise from

alternative RNA processing. It seems unlikely that the difference in transcription could entirely account for the gene conversion defect (Fig. 3.6*A* and *B*). We believe the result reflects the specificity of Region A to target AID function. Although the reporter is controlled by the combination of Region A and Igλ promoter, the AID targeting function is most likely attributable to Region A, since Igλ promoter is transcriptionally inactive by itself (X_p, Fig. 4.2*A*) and is interchangeable with both non-Ig promoters at the Igλ locus (Fig. 3.3*A* and *B*). The gene conversion frequencies of these ectopic substrates are lower relative to that in the Igλ locus (compare X_λ and Y_λ in Fig. 3.6*B* with L1 in Fig. 3.3*B*). The difference may reflect the presence of additional regulatory elements in the Igλ locus as well as position effects on the ectopic substrates.

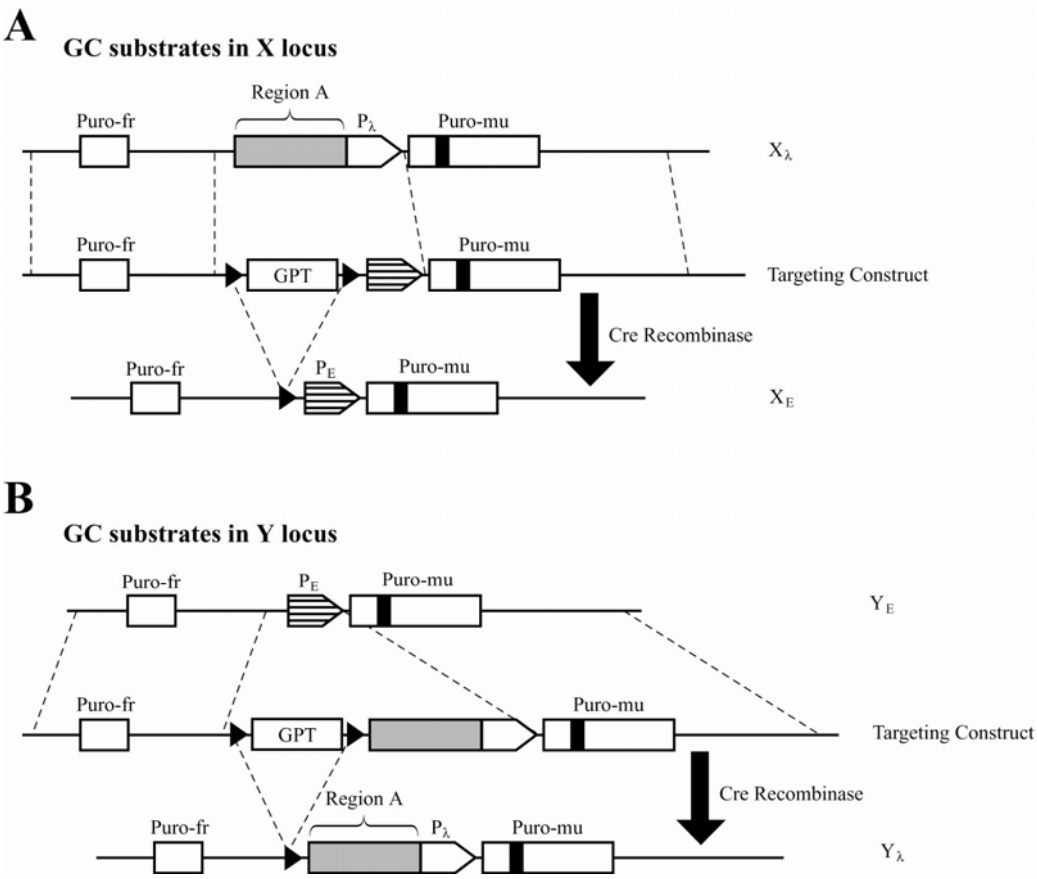

Figure 3.5. Gene conversion substrates in non-Ig locus, X and Y

(A) This diagram illustrates the gene conversion substrates that were randomly integrated into DT40 cells. The integrated locus is named X. Schematic diagram demonstrates replacement of Region A plus Igλ promoter with chicken EF-1α promoter by homologous recombination. (B) The GC substrate with EF-1α promoter without Region A was randomly targeted to the genome of DT40 cells. Schematic diagram illustrates replacement of EF-1α promoter with Region A and Igλ promoter by homologous recombination.

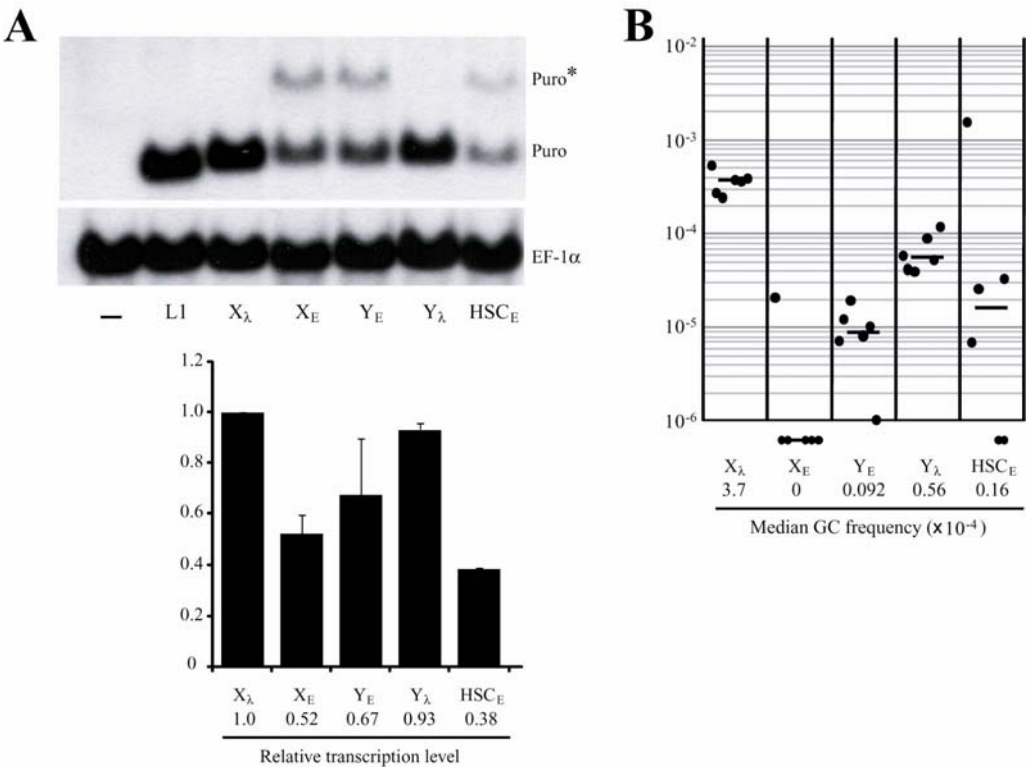

Figure 3.6. Gene conversion analysis at ectopic Loci

(A) This figure shows the transcription levels of ectopic gene conversion substrates. This autoradiograph and histogram are labeled the same way as in Fig. 3.3*A* except that the transcript level of X_λ is set at 1.0 in the histogram. Puro* indicates an aberrant *puro* transcript. (B) The GC frequencies of the substrates, which are randomly integrated and targeted to HSC locus, are shown in this plot. Each dot represents the result from one experiment. The median value is indicated with a bar in the plot and written below the plot.

Substrate	Puro[a]	Control[b]	GC[c]
X_λ	45	42	5.4×10^{-4}
	39	72	2.7×10^{-4}
	32	66	2.4×10^{-4}
	51	68	3.8×10^{-4}
	33	46	3.6×10^{-4}
	48	62	3.9×10^{-4}
X_E	3	73	2.1×10^{-5}
	0	52	0
	0	55	0
	0	63	0
	0	52	0
	0	41	0
Y_E	4	56	7.0×10^{-6}
	7	59	1.2×10^{-5}
	12	62	1.9×10^{-5}
	4	49	8.0×10^{-6}
	6	59	1.0×10^{-5}
	1	70	1.0×10^{-6}
Y_λ	37	63	5.9×10^{-5}
	24	57	4.2×10^{-5}
	25	62	4.0×10^{-5}
	75	83	9.0×10^{-5}
	38	72	5.3×10^{-5}
	107	91	1.2×10^{-4}
HSC_E	192	63	1.5×10^{-3}
	1	72	6.9×10^{-6}
	3	58	2.6×10^{-5}
	0	76	0
	0	69	0
	4	61	3.3×10^{-5}

Table 2. Gene conversion of ectopic substrates.

The data in each row represent analysis of one colony.
a. Puromycin resistant colony per $2x10^5$ cells for X_λ, X_E; per $1x10^6$ cells for Y_λ and Y_E
b. Colony number per 100 cells in media without puromycin
c. Gene conversion frequency = $(a/2x10^5)/(b/100)$ for X_λ, X_E; $(a/10^6)/(b/100)$ for Y_λ and Y_E

3.3. DISCUSSION

We successfully established a novel gene conversion assay system in DT40 cells. The GC assay is more sensitive and less time consuming than other systems that used for GC activity assay with flow cytometry (Arakawa et al., 2002; Sale et al., 2001). More importantly, our GC assay system can be used outside of the endogenous Ig loci. By systematic deletional analysis of the Igλ 3' region, we found that 2.4-kb (Region A) and downstream 7.4-kb (Region B) region can induce gene conversion, indicating that the important elements for AID mediated gene conversion are dispersed throughout 10-kb 3' region. However, since Region A deletion showed more severe reduction of GC activity while maintaining transcription level, we continued to study the roles of Igλ 3' region on targeting gene conversion with Region A. Using the GC assay, we showed that a 2.4-kb regulatory element (Region A) stimulates AID-mediated gene conversion both in and outside of the Igλ locus. The regulatory element contains multiple potential binding sites for transcription factors such as E2A, which has been implicated in AID function (Michael et al., 2003); Schoetz et al., 2006; Conlon and Meyer, 2006; Kitao et al., 2008). It remains to be determined how they collaborate to facilitate AID function. Region A overlaps with *cis*-acting elements characterized in other studies. Deletion of 3'RR impairs both the transcription and gene conversion of Igλ locus (Kothapalli et al., 2008). In a separate report, multiple sequences in the Igλ locus were found to contribute to the diversification of Igλ, and the entire 9.8-kb region from the transcription start site of Igλ to the next non-Ig transcription unit was designated as diversification activator (DIVAC) (Blagodatski et al., 2009). Integration of DIVAS targets AID-mediated hypermutation to non-Ig loci. The overall picture is that multiple regulatory sequences are distributed in the Igλ locus, and Region A characterized in the present study represents one of the AID targeting elements.

CHAPTER 4: Detailed Characterization of Region A [1]

4.1. INTRODUCTION

In chapter 3, we showed that the 2.4-kb Igλ 3' region (Region A) is important for AID mediated gene conversion and is sufficient for targeting gene conversion outside of the Igλ locus. This result is supported by two recent reports demonstrating that Igλ 3' region is critical for gene conversion targeting in DT40 cells (Kothapalli et al., 2008; Blagodatski et al., 2009). However, it is not clear which elements within Region A and what *trans*-acting factors are involved in GC targeting activity.

In order to determine critical elements for GC activity, we analyzed Region A with the GC substrate in ectopic locus, X. We performed detailed deletional analysis of Region A and found that 1.2-kb region (Region A1+A2) can induce transcription and gene conversion activity as efficient as 2.4-kb, Region A. In addition, sequence analysis showed that there are binding sites for nuclear factor-κB (NF-κB) (Combriato and Klobeck, 2001), myocyte enhancer factor-2 (Mef2) (Satyaraj and Storb, 1998), and octamer binding proteins (Oct) within Region A. Mutation analysis of these binding sites within Region A showed notable GC activity reduction with mild transcription activity impairment. As a complementary experiment, we inactivated genes for c-Rel and p50, which are members of NF-κB transcription factor family. Ablation of c-Rel and p50 impairs the GC targeting function of Region A, suggesting that NF-κB family of transcription factors are implicated in AID mediated gene conversion.

---

[1] *Reproduced in part from Kim and Tian (2009) Mol Immunol. (46) 3283-3291. Appropriate permission has been granted by the publisher.*

4.2. RESULTS

4.2.1. Deletional analysis of Region A

As shown previously, Region A is important for targeting GC in the endogenous context, but the ablation of Region A could not abolish GC activity completely. This is because the effects of Region A deletion could be masked by redundant elements dispersed throughout 3' enhancer. However, Region A is sufficient to target gene conversion activity outside of Ig locus, which is a good model to identify enhancer function. To further locate the critical region, we performed systematic deletion analysis with the GC substrate in X locus. To this end, we first narrowed down region A to 442-bp and 727-bp enhancer fragments, designated as Region A1 and Region A2, respectively (X_{A1} and X_{A2}, Fig. 4.1*A*). Region A1 has been known as a core enhancer region among the 10-kb chicken Igλ 3' region (Reynaud et al., 1983; Lauster et al., 1993). Region A2 located downstream of Region A1 contains an octamer binding site and E-boxes. We replaced Region A in the X locus with Region A1 and Region A2 by homologous recombination and found that Region A1 or Region A2 alone cannot induce gene conversion although the GC substrate is transcribed (Fig. 4.1*A* and *B* and Table 3). Next, we replaced Region A with Region A1 combined with Region A2 (X_{A1+A2}, Fig. 4.1*A*). Region A2 plus Region A1 gave rise to GC frequency similar to Region A (Fig. 4.1*B*). The transcriptional activity of Region A1 plus A2 was additive (Fig. 4.1*C*). Therefore, we conclude that Region A1 and A2 contain essential elements for inducing gene conversion, but those elements are located separately in A1 and A2.

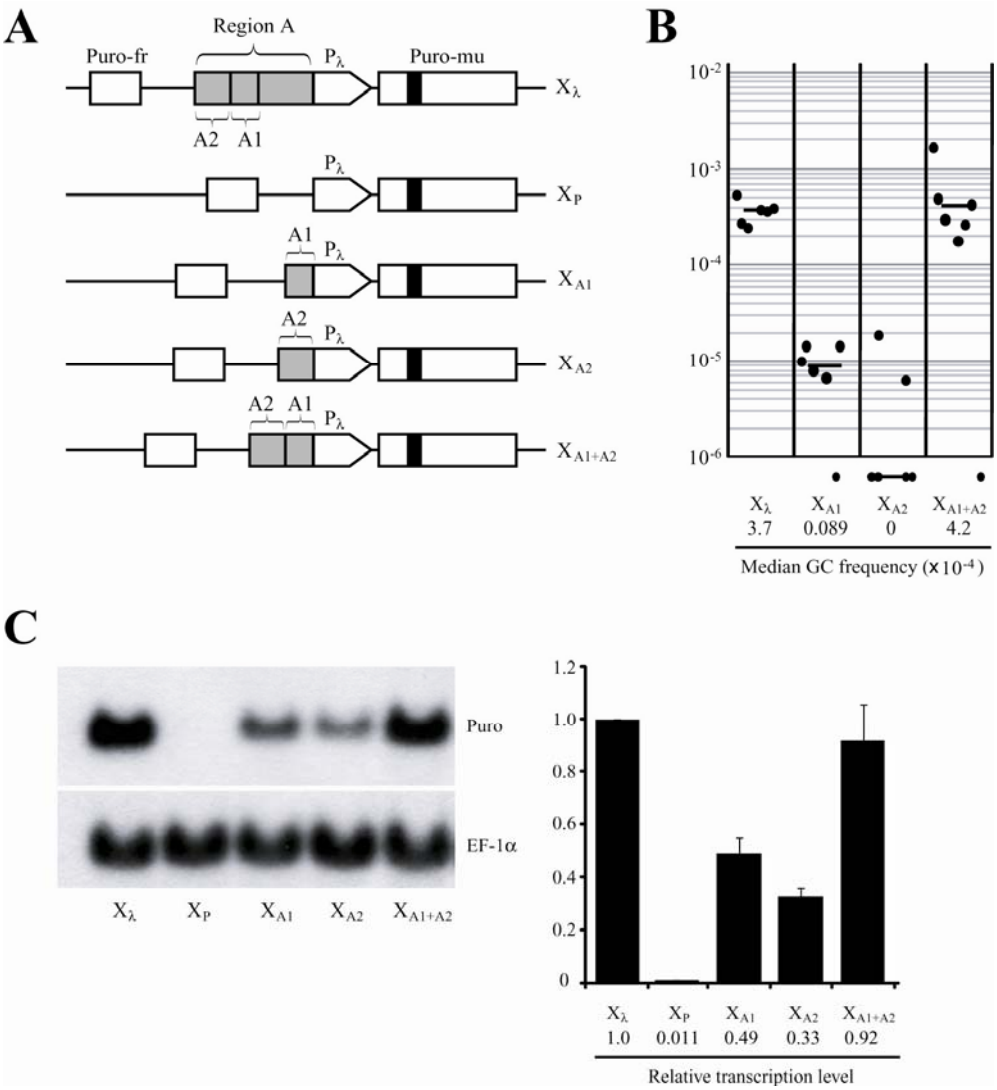

Figure 4.1. Characterization of Region A by deletional analysis

(A) This diagram illustrates the gene conversion substrates that contain various Igλ 3' regions. (B) The GC frequencies of the GC substrates with various Igλ 3' regions are shown in this plot. Each dot represents the result from one experiment. The median value is indicated with a bar in the plot and written below the plot. (C) This figure shows the transcription levels of gene conversion substrates. This autoradiograph and histogram are labeled the same way as in Fig. 3.3. except that the transcript level of X_λ is set at 1.0 in the histogram.

Substrate	Puro[a]	Control[b]	GC[c]
X_λ	45	42	5.4×10^{-4}
	39	72	2.7×10^{-4}
	32	66	2.4×10^{-4}
	51	68	3.8×10^{-4}
	33	46	3.6×10^{-4}
	48	62	3.9×10^{-4}
X_{A1+A2}	183	55	1.7×10^{-3}
	55	56	4.9×10^{-4}
	32	54	3.0×10^{-4}
	29	56	2.6×10^{-4}
	16	45	1.8×10^{-4}
	55	66	4.2×10^{-4}
	50	59	4.2×10^{-4}
X_{A1}	1	51	9.8×10^{-6}
	2	70	1.4×10^{-5}
	1	63	7.9×10^{-6}
	1	76	6.6×10^{-6}
	0	71	0
	2	70	1.4×10^{-5}
X_{A2}	0	40	0
	0	69	0
	2	54	1.9×10^{-5}
	0	69	0
	0	60	0
	1	81	6.2×10^{-6}

Table 3. Gene conversion of ectopic substrates with various Igλ 3' regions

The data in each row represent analysis of one colony.
a. Puromycin resistant colony per $2x10^5$ cells for X_λ, X_{A1}, X_{A2} and X_{A1+A2}
b. Colony number per 100 cells in media without puromycin
c. Gene conversion frequency = $(a/2x10^5)/(b/100)$ for X_λ, X_{A1}, X_{A2} and X_{A1+A2}

4.2.2. Identification of important *cis*-acting elements within Region A

We performed sequence analysis of Region A and found that Region A1 and A2 contain multiple binding sites for transcription factors, including E2A, NF-κB, Mef-2, PU.1, IRF-4 and Oct proteins (Michael et al., 2003; Combriato and Klobeck, 2001; Satyaraj and Storb, 1998; Klein et al., 2006). E2A has been shown to be important for Ig gene conversion and somatic hypermutation (Stoetz et al., 2006). The role of the other factors is unknown. To address this issue, we mutated the binding sites for NF-κB, Mef-2, and Oct protein in X_λ (Fig. 4.2). We found that mutation of each of these binding sites led to approximately 6-fold reduction of AID mediated gene conversion in the context of the ectopic substrate X_λ, while the transcription level of the substrate is largely unaffected (Fig. 4.3*A* and *B* and Table 4). Since the mutations were introduced through homologous recombination, we need to rule out the possibility that the genetic manipulation, rather than the mutation *per se*, underlies those phenotypes. As a control, we performed analogous homologous replacement without introducing the mutation in these binding sites. As shown in figure 4.3 and Table 4 (C1 and C2), such manipulation has no impact on both transcription and gene conversion activity. These experiments suggest that these binding sites, and by implication NF-κB, Mef-2 and Oct proteins, are involved in AID mediated gene conversion.

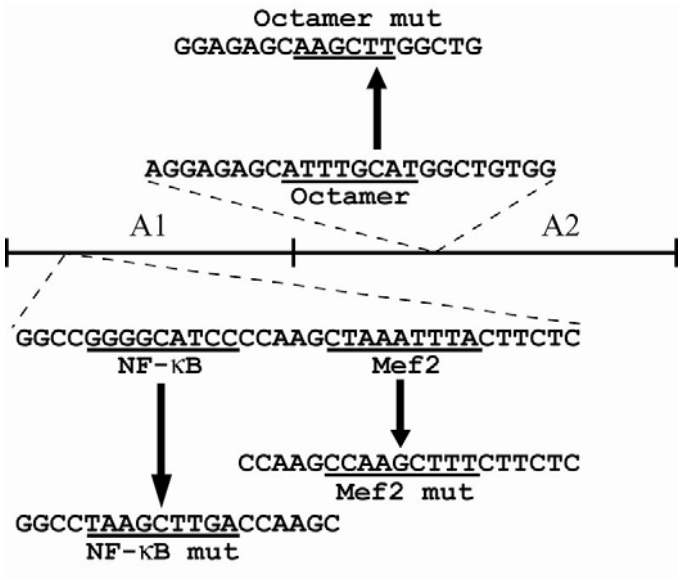

Figure 4.2. Transcription factor binding sites within Region A.

This diagram shows the sequence of the original as well as the mutated binding sites for NF-κB, Mef2 and octamer within A1 and A2 regions.

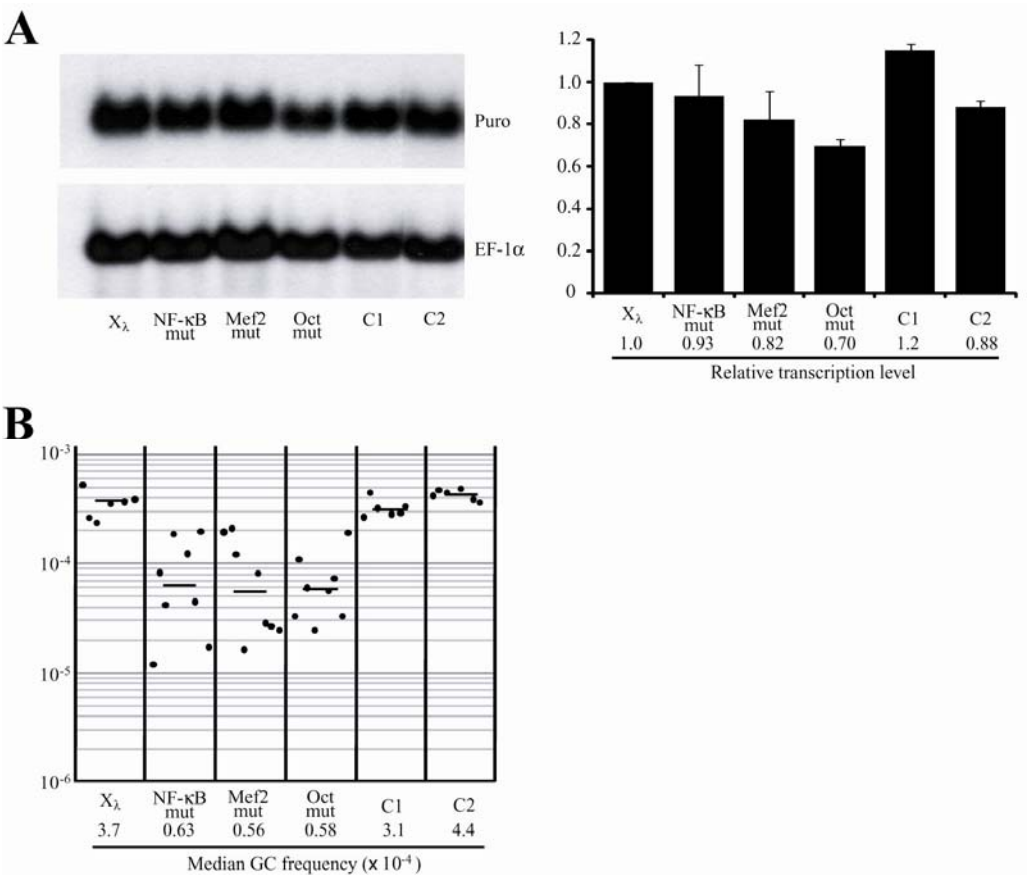

Figure 4.3. Mutation analysis of transcription factor binding sites within Region A

(A) This diagram shows the transcription level of the substrate containing various mutations. The Northern blot and histogram are labeled the same way as Fig. 3.3*A*. (B) This diagram shows the gene conversion frequencies of the constructs containing mutations in the transcription factor binding sites, and C1 and C2 are control replacements without introducing the binding site mutations. The data for wild-type control (X_λ) is the same as the one shown in Fig. 4.1*B*. Abbreviation: mut, mutation.

Substrate	Puro[a]	Control[b]	GC[c]
NF-κB mut	1	41	1.2×10^{-5}
	10	61	8.2×10^{-5}
	5	59	4.2×10^{-5}
	28	74	1.9×10^{-4}
	10	41	1.2×10^{-4}
	4	45	4.4×10^{-5}
	20	50	2.0×10^{-4}
	2	59	1.7×10^{-5}
Mef2 mut	27	70	1.9×10^{-4}
	18	43	2.1×10^{-4}
	16	65	1.2×10^{-4}
	2	62	1.6×10^{-5}
	11	66	8.3×10^{-5}
	4	71	2.8×10^{-5}
	3	55	2.7×10^{-5}
	4	80	2.5×10^{-5}
Oct mut	3	45	3.3×10^{-5}
	12	54	1.1×10^{-5}
	7	59	5.9×10^{-5}
	2	40	2.5×10^{-5}
	7	61	5.7×10^{-5}
	10	67	7.5×10^{-5}
	4	60	3.3×10^{-5}
	16	41	2.0×10^{-4}

Substrate	Puro[a]	Control[b]	GC[c]
C1	25	47	2.7×10^{-4}
	32	35	4.6×10^{-4}
	27	42	3.2×10^{-4}
	39	64	3.0×10^{-4}
	25	41	3.0×10^{-4}
	36	53	3.4×10^{-4}
C2	40	48	4.2×10^{-4}
	56	58	4.8×10^{-4}
	58	63	4.6×10^{-4}
	75	76	4.9×10^{-4}
	55	71	3.9×10^{-4}
	59	79	3.7×10^{-4}

Table 4. Gene conversion of ectopic substrates with mutations in the binding site for NF-κB, Mef2 and octamer proteins

The data in each row represent analysis of one colony.
a. Puromycin resistant colony per $2x10^5$ cells
b. Colony number per 100 cells in media without puromycin
c. Gene conversion frequency = $(a/2x10^5)/(b/100)$

4.2.3. NF-κB family of transcription factors contributes to AID-mediated gene conversion

To directly address the role of NF-κB family of transcription factor, we inactivated the genes for c-Rel and p50 (Fig. 4.4 and 4.5, respectively). For c-Rel knockout, we used two different targeting constructs to inactivate both alleles of c-Rel gene (Fig. 4.4*A*). This strategy is based on our experience that the second targeting event tends to take place on the mutated allele if the same targeting construct is used. In this approach, one of the homology arm of the second targeting construct falls within the region deleted by the first targeting event. Therefore, the second targeting event is restricted to the wild-type allele. Both targeting events delete part of the Rel-homology domain (Kontgen et al., 1995; Pereira and Oakley, 2007), which mediates DNA binding. Truncated c-Rel transcripts were identified in the knockout cell lines with Northern and RT-PCR analysis (Fig. 4.4*B* and *C*). The RT-PCR products were sequenced (data not shown). In the RT-PCR product from the first targeted allele, exon 1 was spliced to exon 5. The joining of exon 1 to exon 5 is in-frame, so the transcript could potentially encode a truncated c-Rel protein without a DNA binding domain. In the RT-PCR product from the second allele, exon 1 is spliced to exon 4, which leads to a frame-shift mutation. Therefore, the homozygous knockout cells should be devoid of functional c-Rel protein.

We took a similar approach to inactivate the p50 genes. We generated two targeting constructs to knockout the two alleles of p50 genes (Fig. 4.5*A*). The first targeting construct deletes exons 6 through 8 of p50 gene, while the second targeting construct eliminates exons 6 and 7. Both mutations disrupt the DNA binding domain (Sha et al., 1995; Beinke and Ley, 2004). We detected truncated p50 transcripts with both Northern and RT-PCR analysis (Fig. 4.5*B* and *C*). We determined the sequences

of the RT-PCR products of truncated transcripts (data not shown). The transcript from the first targeted allele results from the splicing of exon 5 to exon 9, while the transcript from the second targeted allele results from the splicing of exon 5 to exon 8. Both splicing events lead to frame-shift mutation. Thus, the homozygous cells should be null mutant of p50.

Deficiency in p50 or c-Rel does not affect the proliferation and survival of DT40 cells (data not shown). AID transcription is normal as well (Fig. 4.4*B* and 4.5*B*). To test the role of these factors in AID function, we integrated the gene conversion substrate into the Igλ locus of c-Rel and p50 deficient cells (L1, Fig. 3.2*B*). The transcription of the substrate is not affected by the deficiency of either c-Rel or p50 (Fig. 4.6*A*). Gene conversion frequency was reduced by approximately 50% (Fig. 4.6*B* and Table 5). The modest effect of p50 or c-Rel knockout is not surprising since the enhancer containing the NF-κB binding sites is largely dispensable at the endogenous Igλ locus. However, the NF-κB binding site becomes functionally relevant in the ectopic substrate, which is under the control of Region A. Thus, the role of p50 and c-Rel may be revealed when gene conversion is dependent on Region A (Fig. 4.3*B*). To test this possibility, we deleted the 7.3-kb region (Region B, Fig. 3.2*A*) downstream of Region A in the Igλ locus. As assayed with the gene conversion substrate integrated at the Igλ locus (L3, Fig. 3.3*B*), the deletion reduces gene conversion by 5-fold (compare L3/wt with L1/wt in Fig. 4.6*B* and Table 5). In the context of p50 or c-Rel deficiency, the deletion resulted in a more severe defect in gene conversion. Relative to deletion of Region B alone, p50 deficiency reduces gene conversion by 31 fold (compare L3/$p50^{-/-}$ with L3/wt in Fig. 4.6*B*), while c-Rel knockout decreases gene conversion by 6-fold (compare L3/c-$Rel^{-/-}$ with L3/wt in Fig. 4.6*B*). The less severe phenotype of c-Rel deficiency could reflect compensation by redundant NF-κB family members such as p65/RelA. In conjunction

with the data of NF-κB binding site mutation (Fig. 4.3*B*), these results suggest that NF-κB family of transcription factors contribute to AID mediated gene conversion.

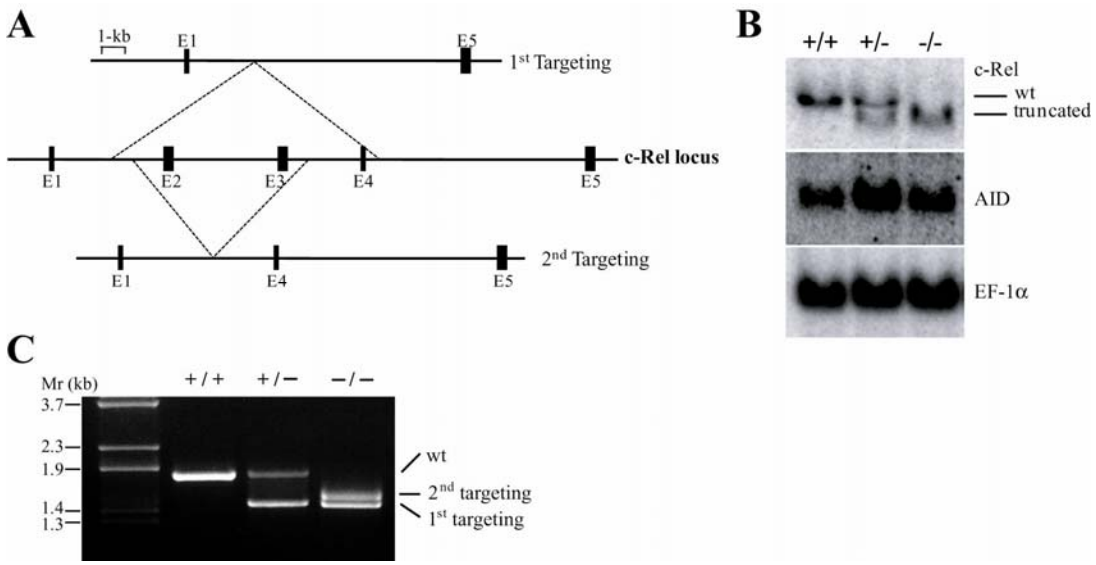

Figure 4.4. Knockout of c-Rel

(A) This diagram illustrates the region of c-Rel gene deleted by the two targeting events. (B) This figure shows the Northern analysis of c-Rel$^{+/+}$, c-Rel$^{+/-}$ and c-Rel$^{-/-}$ cell lines. The probes are indicated next to the autoradiograph. (C) This figure shows the RT-PCR analysis of c-Rel$^{+/+}$, c-Rel$^{+/-}$ and c-Rel$^{-/-}$ cell lines. The PCR product includes the complete open reading frame of c-Rel.

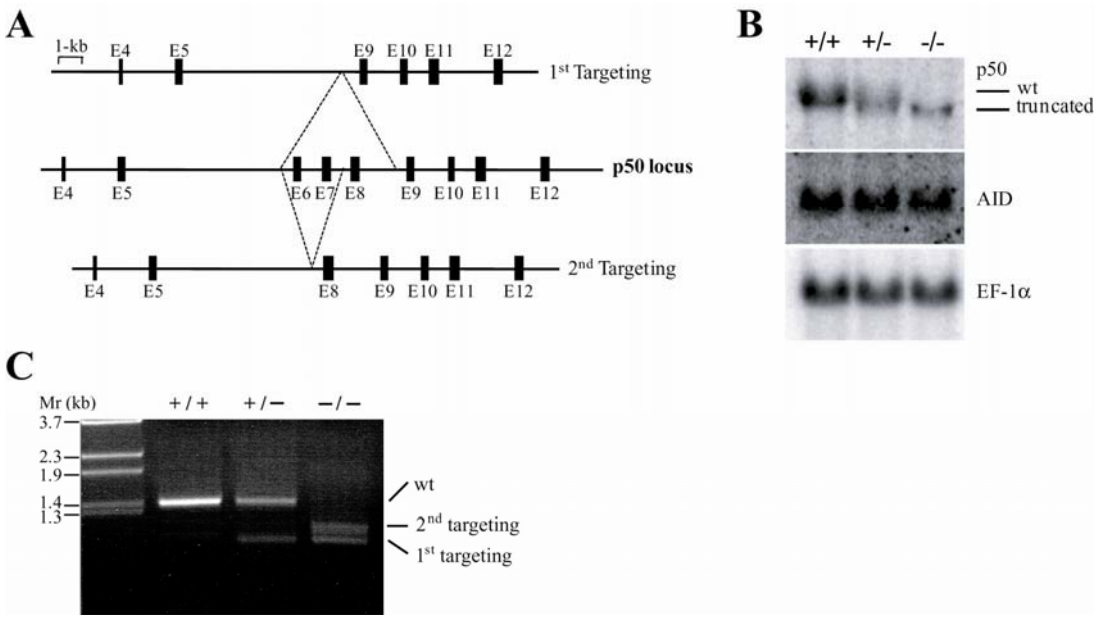

Figure 4.5. Knockout of p50

(A) This diagram illustrates the region of p50 gene that is deleted by the two targeting events. (B) This figure shows the Northern analysis of $p50^{+/+}$, $p50^{+/-}$ and $p50^{-/-}$ cell lines. The probes are indicated next to the autoradiograph. (C) This figure shows the RT-PCR analysis of $p50^{+/+}$, $p50^{+/-}$ and $p50^{-/-}$ cell lines. The PCR product corresponds to the 5' half of p50 open reading frame.

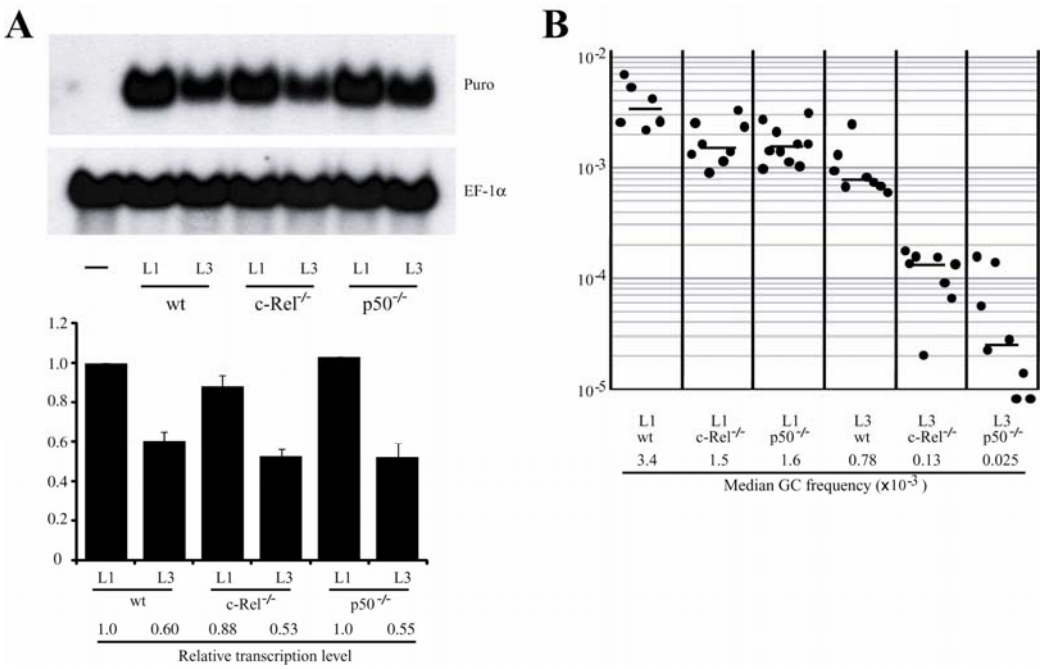

Figure 4.6. Gene conversion analysis of p50 and c-Rel deficient cells

(A) This diagram shows the Northern analysis of the transcription levels of gene conversion substrate in wild-type, c-Rel$^{-/-}$ and p50$^{-/-}$ cells. The Northern blot and histograms are labeled as in figure 3.3*A*. (B) This diagram shows the gene conversion frequency of wild-type, c-Rel$^{-/-}$ and p50$^{-/-}$ cells. The data for L1 is the same as that shown in Fig. 3.3*B*.

Substrate	Puro[a]	Control[b]	GC[c]
	10	38	1.3×10^{-3}
	25	49	2.6×10^{-3}
	18	55	1.6×10^{-3}
L1	9	50	9.0×10^{-4}
c-Rel$^{-/-}$	14	61	1.1×10^{-3}
	12	43	1.4×10^{-3}
	43	65	3.3×10^{-3}
	32	68	2.4×10^{-3}
	30	55	2.7×10^{-3}
	13	44	1.5×10^{-3}
	9	46	9.8×10^{-4}
	19	45	2.1×10^{-3}
L1	19	65	1.5×10^{-3}
p50$^{-/-}$	13	58	1.1×10^{-3}
	14	43	1.6×10^{-3}
	16	49	1.6×10^{-3}
	27	43	3.1×10^{-3}
	14	68	1.0×10^{-3}

Substrate	Puro[a]	Control[b]	GC[c]
	56	60	9.3×10^{-4}
	93	71	1.3×10^{-3}
	39	58	6.7×10^{-4}
L3	151	61	2.5×10^{-3}
wt	37	45	8.2×10^{-4}
	63	86	7.3×10^{-4}
	52	76	6.8×10^{-4}
	39	65	6.0×10^{-4}
	8	46	1.7×10^{-4}
	6	45	1.3×10^{-4}
	10	65	1.5×10^{-4}
L3	1	50	2.0×10^{-5}
c-Rel$^{-/-}$	7	46	1.5×10^{-4}
	5	56	8.9×10^{-5}
	4	62	6.5×10^{-5}
	7	53	1.3×10^{-4}
	8	52	1.5×10^{-4}
	2	36	5.6×10^{-5}
	1	45	2.2×10^{-5}
L3	10	73	1.4×10^{-4}
p50$^{-/-}$	1	36	2.8×10^{-5}
	0	34	0
	1	72	1.4×10^{-5}
	0	61	0

Table 5. Gene conversion at Igλ locus of c-Rel and p50 deficient clones.

The data in each row represent analysis of one colony.

a. Puromycin resistant colony per $2x10^4$ cells for L1 c-Rel$^{-/-}$ and L1 p50$^{-/-}$; per $1x10^5$ cells for L3 wt, L3 c-Rel$^{-/-}$ and L3 p50$^{-/-}$

b. Colony number per 100 cells in media without puromycin

c. Gene conversion frequency = $(a/2x10^4)/(b/100)$ for L1 c-Rel$^{-/-}$ and L1 p50$^{-/-}$; $(a/1x10^5)/(b/100)$ for L3 wt, L3 c-Rel$^{-/-}$ and L3 p50$^{-/-}$

4.3. DISCUSSION

In an effort to narrow down the critical region for AID mediated gene conversion, we found that two synergistic elements (A1 and A2) within the region contribute to its function. Element A1 corresponds to an enhancer that was shown to be dispensable for gene conversion of the Igλ gene, but its function becomes obvious outside of the Ig locus. A 4.1-kb region that contains element A2 together with downstream sequence was defined as 3' regulatory region (3'RR) by another study (Kothapalli et al., 2008). Deletion of 3'RR impairs both the transcription and gene conversion of Igλ locus. In a separate report, multiple sequences in the Igλ locus were found to contribute to the diversification of Igλ locus (Blagodatski et al., 2009). We showed that Region A function involves two synergistic elements, A1 and A2. We analyzed the function of a few binding sites for transcription factors with elements A1 and A2, and found that these binding sites indeed contribute to targeting AID function to gene conversion substrate.

The NF-κB family of transcription factors has long been implicated in the function of immunoglobulin enhancers (Sen, 2004). However, the NF-κB binding site within the mouse Ig intronic enhancer was shown to be dispensable for the transcription and V(D)J recombination of the locus (Inlay et al., 2004), and the function of mouse Igλ 3' enhancers is also independent of NF-κB (Hagman et al., 1990). One potential interpretation of these observations is that the role of NF-κB is masked by redundant transcription factors. Our study supports this view. Ablation of either p50 or c-Rel has minor effect on the gene conversion activity at the endogenous Igλ locus in chicken B cells. The function of NF-κB becomes clear after redundant regulatory elements are removed from the Igλ locus. Similarly, the NF-κB binding site is important for gene conversion in the ectopic substrate, which is dependent on Region A. Besides NF-κB, our study also implicates Mef2 and Octamer binding proteins in the

function of Region A. Mef2 binding site is also present in the mouse Igλ 3' enhancers, and the Mef2 binding site is important for enhancer activity (Satyaraj and Storb, 1998). Several Mef2 family members have been identified, and it is unclear which of the Mef2 family members mediate the function of the Igλ enhancer either in mouse or in chicken (Swanson et al., 1998). The Octamer motif is found in Ig promoters as well as enhancers. The motif can be recognized by either Oct-1 or Oct-2 proteins. In mice, knockout of either Oct-1 or Oct-2 does not affect the expression of immunoglobulin genes, suggesting that the two factors may be functionally redundant in this respect (Wang et al., 2004; Corcoran et al., 1993). In the context of the ectopic substrate, mutation of each binding site (NF-κB, Mef2, and Octamer) causes a similar level of reduction in gene conversion. Our interpretation is that these factors work cooperatively such that loss of one factor impairs the function of the others. None of the binding site mutation has major effect on the transcriptional enhancer activity of the regulatory region. This result implies that their cooperativity lies in targeting AID function, and the transcriptional activity of the regulatory region is mediated by other factors. It remains to be determined how these factors facilitate AID function. A plausible model is that these factors recruit AID or its cofactors to the Ig loci.

CHAPTER 5: Understanding the Mechanism of Region A in Targeting AID function

5.1. INTRODUCTION

Because of the mutagenic function, AID activity needs to be confined to the Ig loci in order to avoid rampant mutations in the genome in general, but the underlying mechanism of targeting AID to Ig loci is largely unknown (Odegard and Schatz, 2006). As we showed earlier, Region A represents an important element for targeting AID function to both the Igλ locus and ectopic sites in DT40 cells; it remains to be established how the targeting function is accomplished by Region A. We propose two hypotheses: first, Region A may increase local chromatin accessibility; second, Region A, through the trans-acting factors associated with it, may recruit AID to adjacent regions. The experiments presented in this chapter are designed to test these hypotheses.

5.2. RESULTS

5.2.1. Region A is not required for maintaining an open chromatin structure at the Ig locus.

It is widely believed that active transcription is associated with local open chromatin structure formations. Since Region A exhibits potent transcriptional enhancer activity (compare X_λ with X_p in Fig. 4.1*C*), it is possible that Region A creates an open chromatin structure and facilitates the access of AID to target DNA. To test this hypothesis, we established a gene conversion system based on the I-SceI endonuclease (Rouet et al., 1994). Since the *puro* reporter contains an I-SceI

recognition sequence, I-SceI should induce DNA double strand break and gene conversion in this substrate. The I-SceI mediated gene conversion frequency ought to correlate with I-SceI cleavage activity. Since susceptibility to nuclease cleavage reflects chromatin accessibility, I-SceI could serve a probe for chromatin structure of the reporter. Relative to DNase I hypersensitivity assay, the advantage of this approach is that it is carried out in live cells, and enables a direct comparison with AID mediated gene conversion. The downside is that gene conversion is an indirect readout of I-SceI cleavage. Although deletion of Region A is not expected to affect homologous recombination in general, we cannot rule out potential local effects.

I-SceI gene conversion system was established in a cell line that expresses the endonuclease in a tetracycline-inducible manner (Fig. 5.1*A*). Gene conversion substrate L1 was integrated at the Igλ locus of this cell line (Fig. 3.2*A*). Since AID gene is ablated in this cell line, gene conversion is dependent on I-SceI, as shown in figure 5.1*B* and table 6 (compare L1(-)dox with L1 (+)dox in Fig. 5.1*B*). Since the *puro* reporter regained the SalI site in all the puromycin resistant colonies (Fig. 5.2), the assay detects I-SceI mediated gene conversion events. To examine the impact of Region A on I-SceI mediated gene conversion, we deleted Region A from Igλ locus (L1 and L2, Fig. 3.2*B*). To our surprise, loss of Region A causes a three-fold increase in gene conversion (compare L2 with L1 in Fig. 5.1*B* and Table 6). To evaluate the generality of this unexpected finding, the same experiment was repeated with substrates with EF-1α and β-actin promoters. Consistently, deletion of Region A increased gene conversion in these substrates as well (L4 through L7, Fig. 5.1*B* and Table 6). The effect is most obvious with β-actin promoter. Intriguingly, deletion of Region A causes the most severe reduction in AID-mediated gene conversion in the context of the β-actin promoter (Fig. 3.3B and Table 1). To look into this issue further, we integrated the gene conversion

substrate into a region upstream of the hsc70 gene. As expected, AID-mediated gene conversion activity is low for this substrate in DT40 cells (HSC_E, Fig. 3.6*B,* and Table 2), even though the reporter is transcribed (HSC_E, Fig. 3.6*A*). By contrast, the gene conversion activity of this substrate in I-SceI inducible cell line is higher than that of the substrate at the Igλ locus with Region A (L1, L4, L6 and HSC in Fig. 5.1*B* and Table 6). These data suggest that Region A has opposite effects on AID versus I-SceI mediated gene conversion. The cause of this phenomenon is presently unclear, but loss of Region A does not appear to decrease chromatin accessibility.

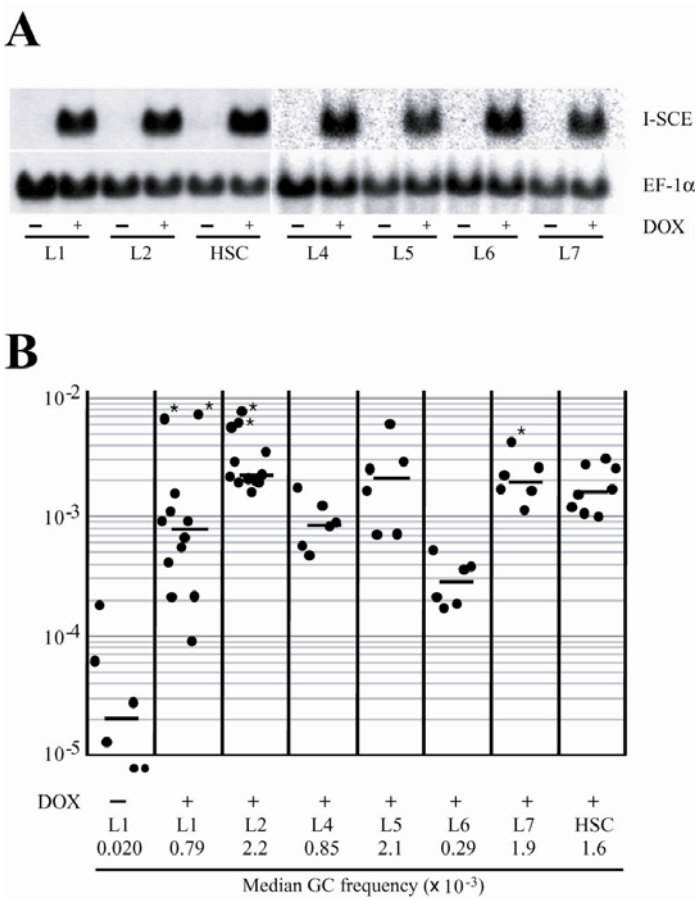

Figure 5.1. Analysis of I-SceI mediated gene conversion

(A) The autoradiographs are Northern analysis of I-SceI expression in different cell lines before and after doxcycline induction. The source of RNA is indicated below the autoradiograph. Each blot was hybridized sequentially with probes for I-SceI and EF-1α. (B) This plot shows the frequencies of I-SceI mediated gene conversion. As indicated in Table 6, more than 200 puromycin resistant colonies were generated in five experiments, and it is difficult to count the colony number with accuracy. These data are represented with an asterisk (*) in the plot. The value was calculated by assuming 200 puromycin resistant colonies, and represents the lower limit of the real gene conversion frequency. The estimation does not affect the median value of the whole data set. L1(-)dox represents gene conversion experiment without induction of I-SceI by doxcycline. All the other experiments were performed in the presence of doxcycline and I-SceI expression.

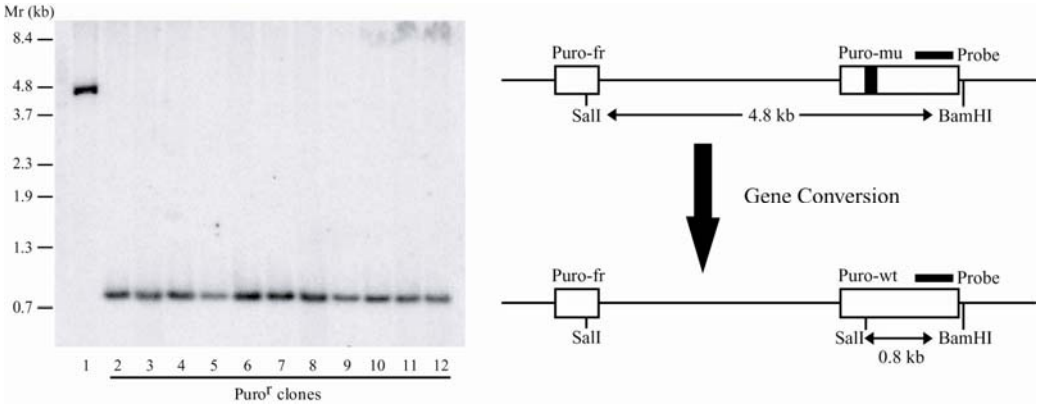

Figure 5.2. Regeneration of SalI site by gene conversion

This autoradiograph is Southern blot analysis of DNAs from total L2 cells (lane 1) or from puromycin resistant clones after I-SceI induction (lane 2 through 12). The DNA was digested with BamHI and SalI. The location of the probe and expected restriction digest fragments are indicated in the diagram right of the autoradiograph.

Substrate	Puro[a]	Control[b]	GC[c]
L1/ISCE -dox	4	65	6.2×10^{-5}
	12	66	1.8×10^{-4}
	1	78	1.3×10^{-5}
	0	65	0
	0	60	0
	2	73	2.7×10^{-5}
L1/ISCE	25	45	1.1×10^{-3}
	35	76	9.2×10^{-4}
	>200[d]	60	$>6.7\times10^{-3}$
	12	58	4.1×10^{-4}
	6	57	2.1×10^{-4}
	50	64	1.6×10^{-3}
	11	40	5.5×10^{-4}
	15	45	6.7×10^{-4}
	22	48	9.2×10^{-4}
	3	66	9.1×10^{-5}
	7	65	2.2×10^{-4}
	>200[d]	55	$>7.3\times10^{-3}$
L2/ISCE	59	61	1.9×10^{-3}
	77	71	2.2×10^{-3}
	165	58	5.7×10^{-3}
	103	71	2.9×10^{-3}
	>200[d]	64	$>6.3\times10^{-3}$
	>200[d]	52	$>7.7\times10^{-3}$
	50	48	2.1×10^{-3}
	51	63	1.6×10^{-3}
	45	45	2.0×10^{-3}
	69	71	1.9×10^{-3}
	64	57	2.2×10^{-3}
	99	56	3.5×10^{-3}

Substrate	Puro[a]	Control[b]	GC[c]
L4/ISCE	50	58	1.7×10^{-3}
	15	53	5.7×10^{-4}
	14	59	4.7×10^{-4}
	37	60	1.2×10^{-3}
	18	44	8.2×10^{-4}
	20	45	8.9×10^{-4}
L5/ISCE	33	40	1.7×10^{-3}
	54	43	2.5×10^{-3}
	20	56	7.1×10^{-4}
	130	43	6.0×10^{-3}
	22	61	7.2×10^{-4}
	81	56	2.9×10^{-3}
L6/ISCE	34	65	5.2×10^{-4}
	14	66	2.1×10^{-4}
	12	70	1.7×10^{-4}
	9	49	1.8×10^{-4}
	34	95	3.6×10^{-4}
	29	76	3.8×10^{-4}
L7/ISCE	95	56	1.7×10^{-3}
	99	45	2.2×10^{-3}
	>200[d]	47	$>4.3\times10^{-3}$
	80	70	1.1×10^{-3}
	68	41	1.7×10^{-3}
	151	59	2.6×10^{-3}
HSC/ISCE	33	62	1.1×10^{-3}
	40	67	1.2×10^{-3}
	48	63	1.5×10^{-3}
	55	40	2.8×10^{-3}
	38	76	1.0×10^{-3}
	99	65	3.0×10^{-3}
	56	66	1.7×10^{-3}
	83	65	2.6×10^{-3}

Table 6. I-SceI mediated gene conversion

The data in each row represent analysis of one colony.

a. Puromycin resistant colony per 10^5 cells for L1/ISCE-dox, L6/ISCE, and L7/ISCE; per $5x10^4$ cells for L1/ISCE, L2/ISCE, L4/ISCE, L5/ISCE, and HSC/ISCE.

b. Colony number pre 100 cells in media without puromyin.

c. Gene conversion frequency = $(a/10^5)/(b/100)$ for L1/ISCE-dox, L6/ISCE, and L7/ISCE; $(a/5x10^4)/(b/100)$ for L1/ISCE, L2/ISCE, L4/ISCE, L5/ISCE, and HSC/ISCE.

d. There are more than 200 puromycin resistant colonies, and it is difficult to count the number with accuracy.

To further address the role of Region A in regulating chromatin accessibility, we used *E. coli* DNA adenine methylatransferase (Dam) as a probe (Buryanov and Shevchuk, 2005), and analyzed the impact of Region A deletion on the accessibility of Igλ locus toward Dam. Dam methylates the adenine residue in the sequence of GATC (Palmer and Marinus, 1994). The methylation status of GATC can be determined by a methylation sensitive restriction endounclase, such as MboI which cleaves un-methylated GATC sequences only. There is no endogenous methylation of adenine in eukaryotic cells (van Steensel and Henikoff, 2000). Expression of Dam in eukaryotic cells led to the methylation of GATC motifs and the methylation efficiencies appear to correlate with local chromatin accessibility (Venetianer et al., 1983). For example, when Dam was expressed in Yeast, it was found that the promoter regions of actively transcribed genes are efficiently methylated (Singh and Klar, 1992). In contrast, telomere-proximal region DNAs are refractory to Dam methylation (Gottschling, 1992; Wright et al., 1992). We applied this method to our experiments and established an inducible Dam expression system in DT40 cells (Fig. 5.3*A* and *B*). We introduced the Dam expression construct into the DT40 cells that contain the gene conversion substrate either with or without Region A (L1 and L2 DT40 clone, respectively). In this construct, the *dam* gene is cloned downstream of the coding region for gpt selection marker, which is flanked with loxP sites. Deletion of the selection marker via cre-mediated recombination removes the selection marker and places the dam gene under the control of the chicken β-actin promoter (Fig. 5.3*A*). This induction system ensures very low levels of basal expression (Fig. 5.3*B*). This feature is critical since our initial attempt with tetracycline inducible system failed because leaky expression of Dam is sufficient to cause extensive methylation of the genome.

We integrated the inducible dam construct into the AID locus of both L1 and L2 cell lines. The integration into the same locus ensures that dam expression is comparable in both cell lines (Fig. 5.3*B*). Tamoxifen was added into the culture to induce the expression of Dam, and genomic DNA was harvested at different time points after Dam induction. Dam methylation activity in DT40 cell line was detected with sensitivity to MboI digest (Fig. 5.3*C*). MboI cleaves only unmethylated GATC sequence. Therefore, resistance to MboI cleavage is an indication of methylation at that GATC site.

To test whether Dam methylation efficiency is affected by chromatin accessibility in DT40 cells, we measured DNA methylation frequencies in four different loci: Two highly transcribed regions and two un-transcribed regions (Fig. 5.4*A* to *D*). Actively transcribed regions generally exist in open chromatin structures, while un-transcribed regions are usually within closed chromatin structures (Singh and Klar, 1992). We determined the methylation status of Igλ locus (Fig. 5.4*A*)and EF-1α locus (Fig. 5.4*B*)which are highly transcribed regions, and AID 3' intergenic region (Fig. 5.4*C*) and β-globin locus (Fig. 5.4*D*) as examples of un-transcribed regions. Dam expression was induced by adding tamoxifen to the cell culture, and genomic DNAs from the L1 DT40 clone, which contains the GC substrate at the Igλ locus, were harvested at 0, 8, and 12 hours after tamoxifen addition. The genomic DNAs were digested with MboI in combination with other enzymes (Fig. 5.4*A* to *D*). The extent of DNA methylation was analyzed by Southern blot and the blots were quantified with Phophoimager. As shown in figure 5.4*E*, Igλ and EF-1α region were methylated more efficiently than AID 3' intergenic and β-globin loci. Therefore, Dam methylation activity appears to correlate with the transcription activity, and presumably chromatin accessibility, of the loci, consistent with observations made in other systems.

Using this assay, we probed the methylation status of two GATC sites within the *puro* gene of the gene conversion substrate in the endogenous Igλ locus. We found that the kinetic of methylation of these two GATC sites is identical with or without the Region A, (Fig. 5.4*F*) even though there is a substantial difference in gene conversion frequencies between the two situations as shown in figure 3.3*B*. Thus, in combination with the results from the I-SceI experiment, these data further reinforce the view that the function of Region A does not lie in creating an open chromatin structure. By comparison with the I-SceI experiment, we did not observe the inhibitory effect of Region A on Dam methylation activities. The cause of this difference is presently unclear.

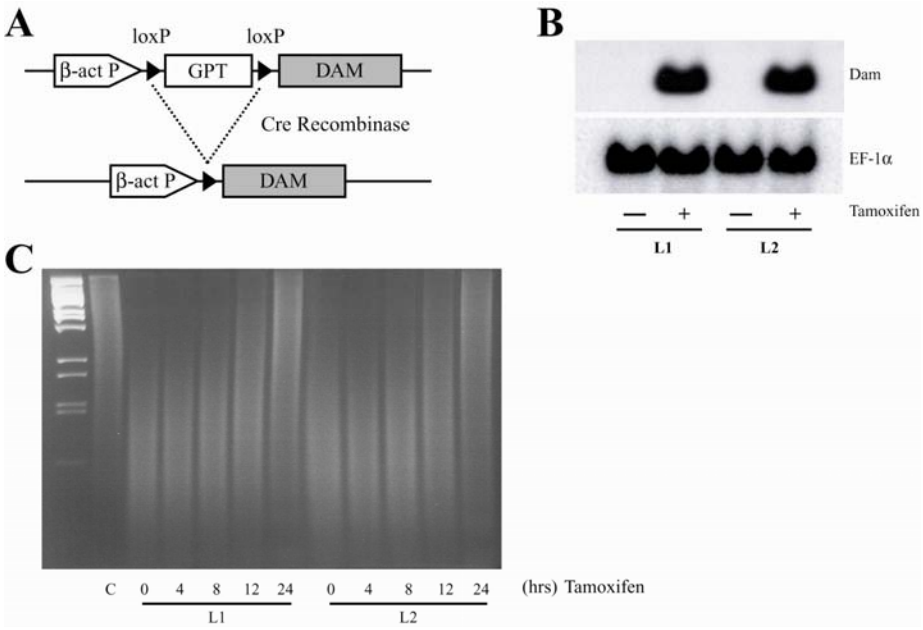

Figure 5.3. Expression of the *E.coli* DNA methyltransferase in DT40 cells

(A) This diagram illustrates the inducible dam expression construct. (B) The autoradiographs are Northern blot analysis of Dam expression in L1 and L2 cell lines before and after tamoxifen treatment. The source of RNA is indicated below the autoradiograph. Each blot was hybridized sequentially with probes for Dam and EF-1α. (C) Genomic DNA stained with ethidium bromide after overnight EcoRI, XbaI and MboI digestion. Genomic DNAs were harvested after 0, 4, 8, 12, and 24 hours of Dam induction. As a control, genomic DNA was digested by EcoRI and XbaI (Lane C).

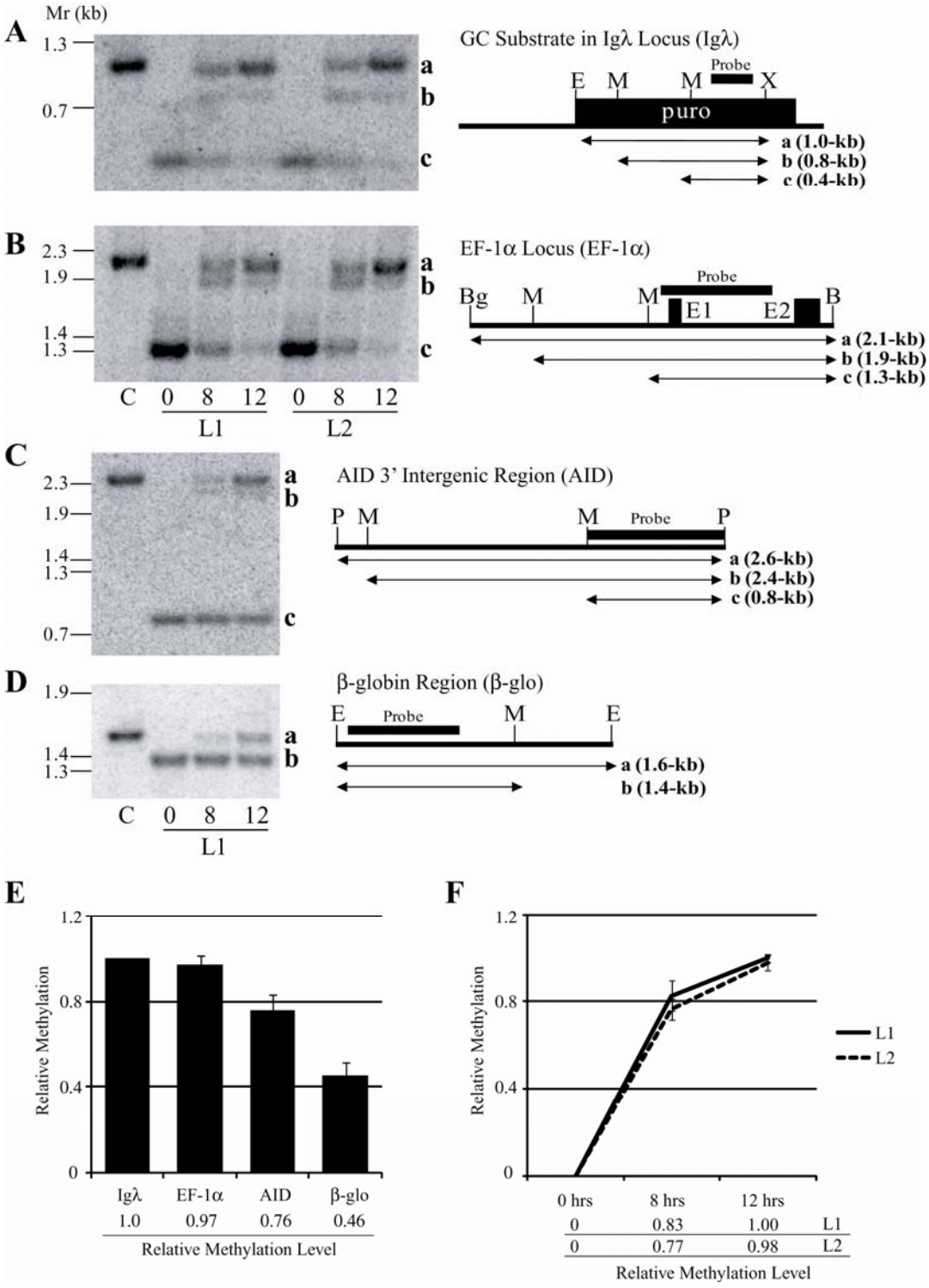
A
Mr (kb)
1.3
0.7
a
b
c
GC Substrate in Igλ Locus (Igλ)
Probe
E M M X
puro
a (1.0-kb)
b (0.8-kb)
c (0.4-kb)
B
2.3
1.9
1.4
1.3
a
b
c
EF-1α Locus (EF-1α)
Probe
Bg M M E1 E2 B
a (2.1-kb)
b (1.9-kb)
c (1.3-kb)
C 0 8 12 0 8 12
L1 L2
C
2.3
1.9
1.4
1.3
0.7
a
b
c
AID 3' Intergenic Region (AID)
P M M Probe P
a (2.6-kb)
b (2.4-kb)
c (0.8-kb)
D
1.9
1.4
1.3
a
b
β-globin Region (β-glo)
E Probe M E
a (1.6-kb)
b (1.4-kb)
C 0 8 12
L1
E
1.2
0.8
0.4
0
Relative Methylation
Igλ EF-1α AID β-glo
1.0 0.97 0.76 0.46
Relative Methylation Level
F
1.2
0.8
0.4
0
Relative Methylation
L1
L2
0 hrs 8 hrs 12 hrs
0 0.83 1.00 L1
0 0.77 0.98 L2
Relative Methylation Level

Figure 5.4. Methylation analysis of the gene conversion substrate

(A) This Southern blot shows the MboI digest pattern of gene conversion substrate in the endogenous Igλ locus. The genomic DNAs from L1 and L2 cells were harvested at 0, 8, and 12 hours of Dam induction, and digested with MboI+EcoRI+XbaI. The DNA in lane C was digested with EcoRI+XbaI. The location of these restriction sites is indicated in the diagram, right side of the autoradiograph. The size of the restriction fragments and the probe are indicated. (B),(C),(D) Southern blot analysis show the MboI digest pattern of EF-1α locus, AID 3' intergenic region and β-globin locus, respectively. Same genomic DNAs from (A) were digested by MboI+BamHI+BglII for EF-1α locus, MboI+PvuII for AID 3' intergenic region, and MboI+EcoRI for β-globin locus. The DNA in lane C was digested with BamHI+BglII for EF-1α locus, PvuII for intergenic region and EcoRI for β-globin locus. The location of all these restriction sites is indicated in the diagram, respectively. (E) This figure shows the relative methylation frequency in Igλ, EF-1α, AID 3' intergenic region, and β-globin locus. The analyzed regions were indicated below the histogram. Methylation level is based on quantification of hybridization signals from Southern blotting with Phosphoimager. Methylation level of Igλ locus is set at 1.0. The histogram represents the average of three experiments, and the average value is written below the graph. Error bars represent standard deviations. (F) This graph illustrates relative methylation of the gene conversion substrates in the endogenous Igλ locus with (L1) and without (L2) Region A. Methylation level of the gene conversion substrates is normalized with that of EF-1α locus. The extent of methylation after 12 hours Dam induction in L1 is set at 1.0. The graph represents the average of three experiments and the average value is indicated below the time points. Error bars represent standard deviations. Abbreviations: E, EcoRI; X, XbaI; M, MboI; Bg, BglII; B, BamHI; P, PvuII; AID, AID 3' intergenic region; β-glo, β-globin

5.2.2. Region A facilitates the recruitment of AID to the Igλ locus.

Chromatin immune-precipitation (ChIP) is a widely used approach to examine protein/DNA interactions *in vivo*. An alternative approach (DamID), based on the *E. coli* Dam methyltransferase, has also been developed for the same purpose (Steensel and Henikoff, 2000). In this method, a protein of interest is fused with Dam, and the fusion protein is expressed in the cell. If the fusion partner binds to a particular DNA sequence in the genome, it would target the associated Dam to the site and cause methylation of the locus. Relative to the ChIP method, the advantage of DamID is that it does not require an antibody against the protein of interest, which could be a limiting factor for ChIP. Moreover, DamID does not involve formaldehyde treatment, which could affect chromatin structure (Steensel and Henikoff, 2000). We adapted this method to examine the interaction of AID with the Igλ locus. We fused AID with Dam, and expressed the fusion protein with the tamoxifen inducible construct (Fig. 5.5*A*). The DNA construct was targeted to the AID locus to inactivate the expression of endogenous AID, and also to ensure uniformed expression of the fusion among different cell lines. We introduced the expression construct into both L1 and L2 DT40 clones. As described before, L1 clone contains the GC substrate integrated into the Igλ locus with intact 3' region, while L2 contain contains a deletion of Region A. As judged with Northern analysis, AID-Dam mRNA was expressed at the same levels in the both cell lines after tamoxifen induction (Fig. 5.5*B*). Before performing the methylation analysis, we first tested the activity of the fusion proteins in gene conversion. In these cells, both alleles of the endogenous AID genes were inactivated so that gene conversion in these cell types would have to be dependent on the AID-Dam fusion protein (Fig. 5.5*A*). Before tamoxifen induction, there is no detectable gene conversion activity in these cell line (Fig. 5.5*C* and Table 7). The addition of tamoxifen led to the expression of AID-Dam and concomitant

gene conversion in L1 cells (L1 (+) AID-Dam expression, Fig. 5.5*C* and Table 7), indicating that the AID-Dam fusion protein is functional for gene conversion. Moreover, the function of the AID-Dam fusion is dependent on Region A since, relative to L1 cells, gene conversion activity in L2 is reduced by 31-fold (L2 (+) AID-Dam expression, Fig. 5.5*C* and Table 7). Thus, the AID-Dam fusion behaves similarly as normal AID proteins in gene conversion assays, and it would be relevant to examine the interaction of this fusion protein with the Igλ locus. After induction of the expression of AID-Dam, we could indeed detect methylation of the GC substrate in both L1 and L2 cells, and found that the *puro* reporter in L1 cells is methylated faster relative to that in L2 cells (Fig. 5.7*A* and *D*). As an internal control for AID-Dam activity, we tested the methylation status of EF-1α locus in L1 and L2 cells, and found little difference in methylation levels of the EF-1α locus between these two cell types (Fig. 5.7*A*). This result indicates that the deaminase activity of AID-Dam fusion protein is comparable in L1 and L2 cells, and the difference in methylation levels at the *puro* reporter could be attributed to differences in local AID concentration or DNA accessibility. As the experiments with free Dam and I-SecI mediated gene conversion suggest that Region A does not appear to affect local chromatin accessibility, we think the difference is primarily attributable to preferential recruitment of AID to the locus in the presence of Region A. To quantify the difference, we measured the methylation levels in both the GC substrate and EF-1α locus by Phosphoimager, and normalized the methylation levels of the GC substrate to that of the EF-1α locus. As shown in figure 5.7*D*, there was about 56% and 40% reduction after 8 and 12 hours of AID-Dam induction, respectively, in the absence of Region A.

To confirm that the role of Region A in AID-Dam recruitment is specific to AID, we constructed the APOBEC1-Dam fusion protein expression vector and transfected it

into L1 and L2 cells (Fig. 5.6*A*). The expression level of APOBEC1-Dam was comparable in L1 and L2 clone upon tamoxifen treatment (Fig. 5.6*B*), and we performed DNA methylation assay in the GC substrate and EF-1α loci in an analogous fashion as with the AID-Dam fusion. As shown in figure 5.7*B*, there was a little difference in methylation by APOBEC1-Dam at the Ig locus with and without Region A. As another control, we generated a PKA-Dam fusion expression construct (Fig. 5.6*A* and *C*), and performed the same DNA methylation assay. Like APOBEC1-Dam, Region A did not change PKA-Dam fusion protein accessibility to Ig locus (Fig. 5.7*C*), indicating that indeed the Region A facilitates the recruitment of AID specifically to Igλ locus. Furthermore, in comparison with AID-Dam, the methylation activities of APOBEC1-Dam and PKA-Dam at the to Igλ locus were low in reference to the EF-1α locus, again suggesting preferential recruitment of AID to the Igl locus (Fig. 5.7*E*).

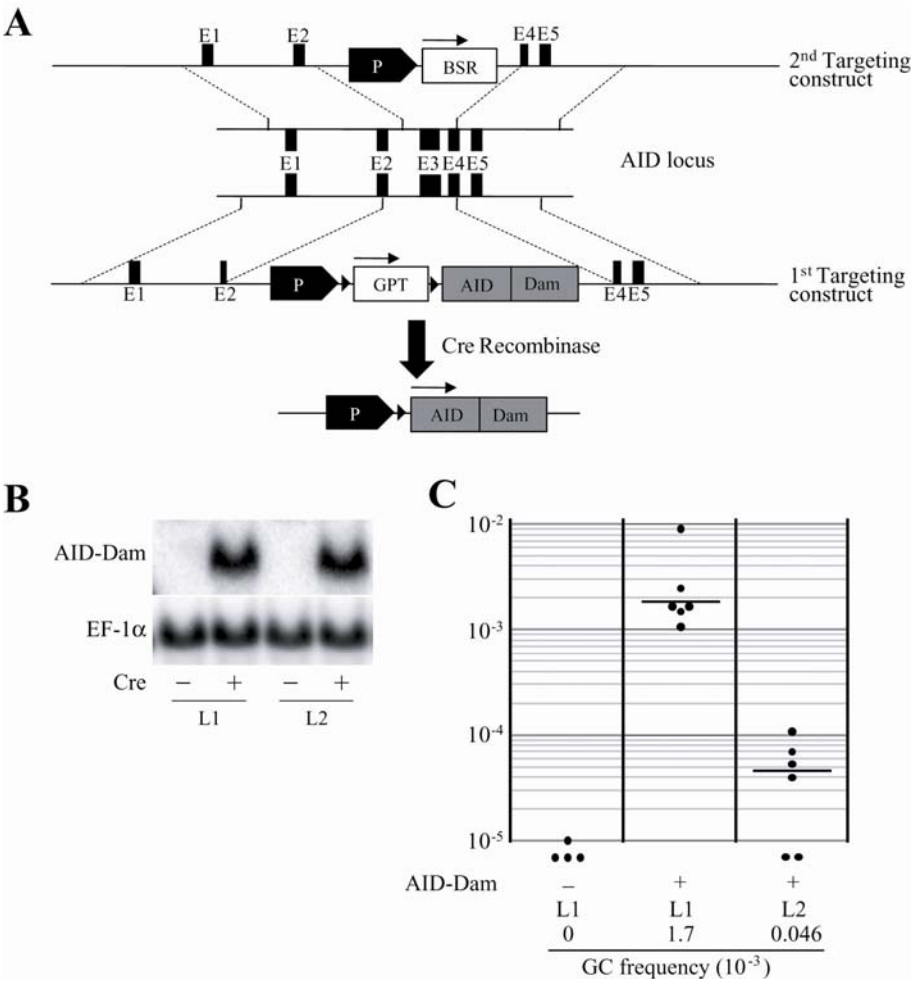

Figure 5.5. Inducible AID-Dam fusion protein expression

(A) This diagram illustrates the inducible AID-Dam fusion protein expression construct, which is integrated to AID locus. The other intact AID allele is disrupted by another round of targeting event. Deleted region is illustrated. (B) This Northern blot analysis shows the induction of AID-Dam fusion protein by tamoxifen treatment. (C) AID-Dam mediated gene conversion frequency of L1 and L2 are shown in this plot. Each dot represents the data from one experiment. The median value is indicated with a bar and written below the plot.

Substrate	Puro[a]	Cont[b]	GC[c]
L1 (-) HTX	0	72	0
	1	79	6.3×10^{-6}
	1	157	3.2×10^{-6}
	0	143	0
L1 (+) HTX	275	61	9.0×10^{-3}
	39	45	1.7×10^{-3}
	67	125	1.1×10^{-3}
	116	135	1.7×10^{-3}
	73	59	2.5×10^{-3}
	106	143	1.5×10^{-3}
L2 (+) HTX	0	127	0
	7	130	1.1×10^{-4}
	3	87	6.9×10^{-5}
	2	76	5.3×10^{-5}
	1	51	3.9×10^{-5}
	0	40	0

Table 7. AID-Dam mediated gene conversion assay

The data in each row represent analysis of one colony.

a. Puromycin resistant colony per 10^6 cells for L1/(-) HTX; per $2.5\text{x}10^5$ cells for L1/(+) HTX and L2/(+) HTX

b. Colony number pre 500 cells in media without puromyin.

c. Gene conversion frequency = $(a/10^6)/(b/500)$ for L1/(-) HTX; $(a/2.5\text{x}10^5)/(b/500)$ for L1/(+) HTX and L2/(+) HTX

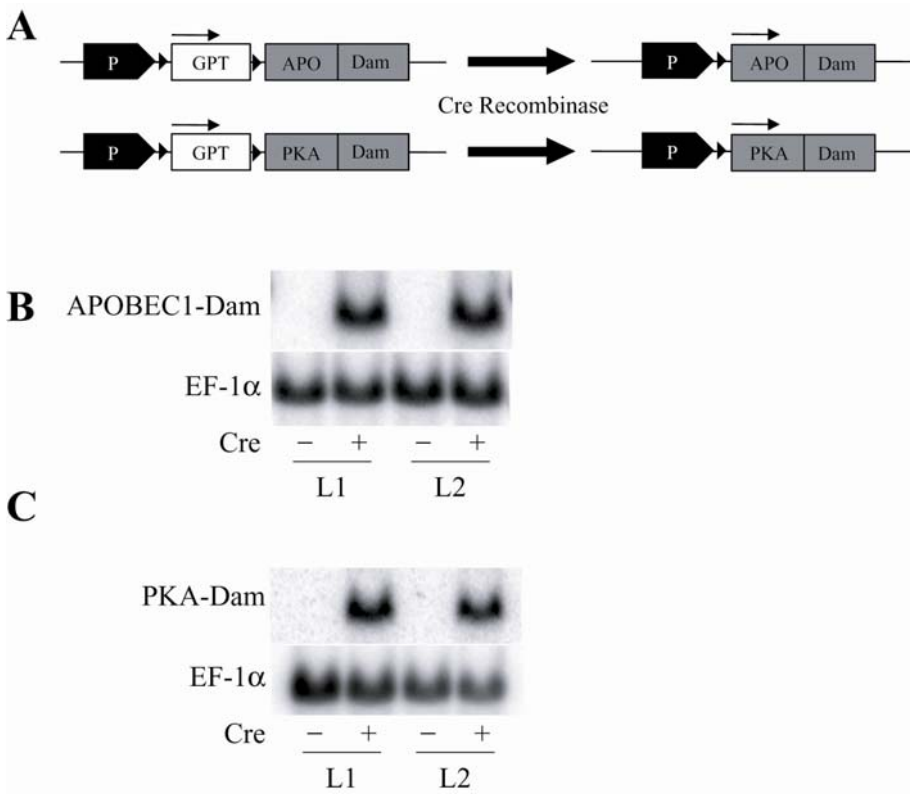

Figure 5.6. Inducible APOBEC1-Dam and PKA-Dam fusion protein expression

(A) These diagrams illustrate tamoxifen dependent APOBEC1-Dam and PKA-Dam fusion protein expression constructs. (B) and (C) Northern blot analyses show the induction of APOBEC1-Dam and PKA-Dam fusion proteins by tamoxifen treatment in L1 and L2 DT40 clones. Abbreviations: APO, APOBEC1; PKA, protein kinase A catalytic subunit.

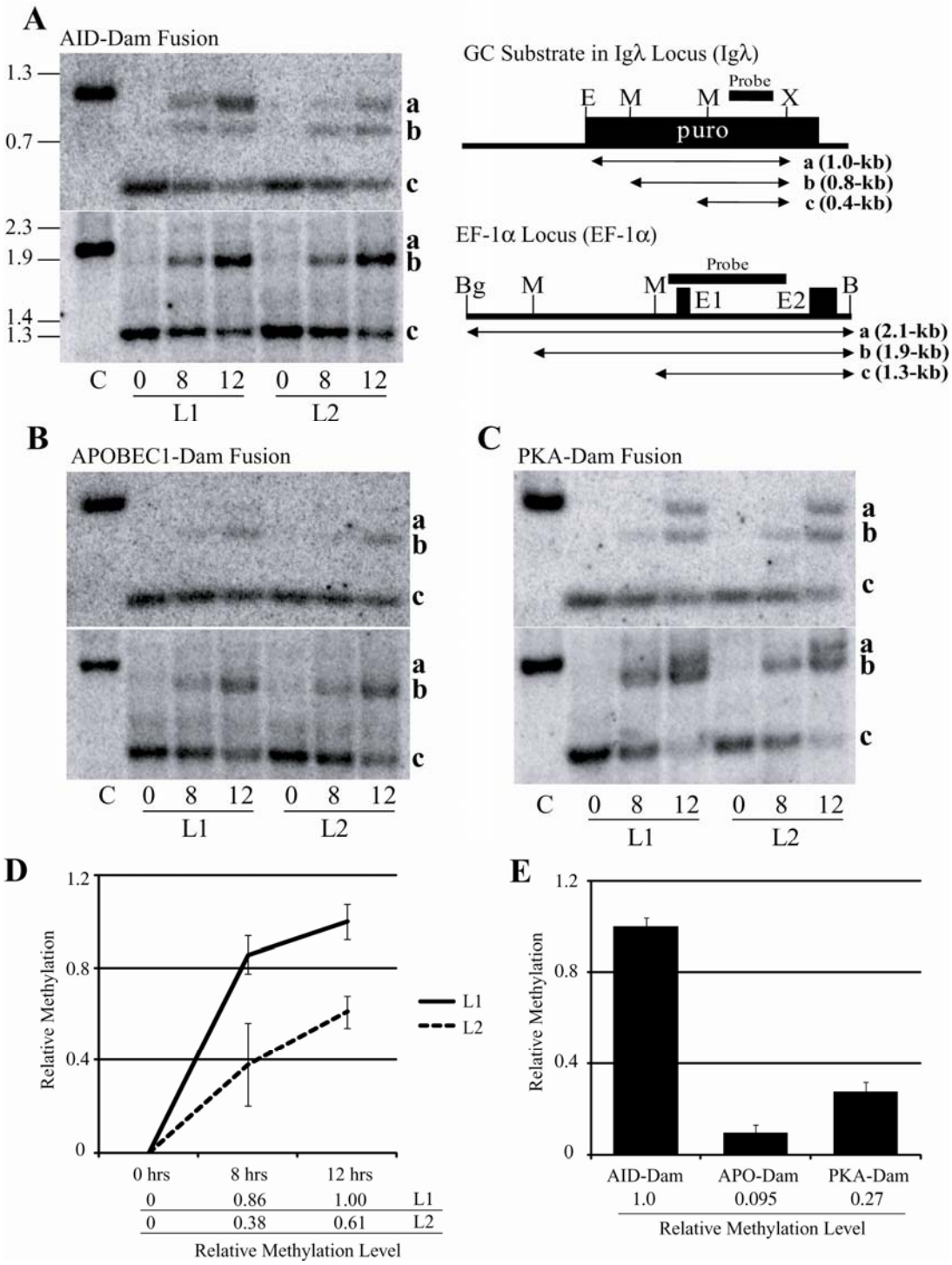
A
AID-Dam Fusion
1.3
0.7
2.3
1.9
1.4
1.3
a
b
c
C 0 8 12 0 8 12
L1 L2
GC Substrate in Igλ Locus (Igλ)
Probe
E M M X
puro
a (1.0-kb)
b (0.8-kb)
c (0.4-kb)
EF-1α Locus (EF-1α)
Probe
Bg M M E1 E2 B
a (2.1-kb)
b (1.9-kb)
c (1.3-kb)
B
APOBEC1-Dam Fusion
C
PKA-Dam Fusion
D
Relative Methylation
1.2
0.8
0.4
0
L1
L2
0 hrs 8 hrs 12 hrs
0 0.86 1.00 L1
0 0.38 0.61 L2
Relative Methylation Level
E
AID-Dam APO-Dam PKA-Dam
1.0 0.095 0.27
Relative Methylation Level

Figure 5.7. Involvement of Region A in AID recruitment to Igλ locus

(A) This Southern blot shows the MboI digest pattern of gene conversion substrate in the endogenous Igλ locus. The genomic DNAs from L1 and L2 cells were harvested after 0, 8, and 12 hours induction of AID-Dam fusion protein and digested by MboI+EcoRI+XbaI. The DNA in lane C was digested with EcoRI+XbaI. The location of these restriction sites is illustrated in the diagram, right side of the autoradiograph. The size of the restriction fragments and the probe are indicated. (B), (C) Same experiment was performed with genomic DNA harvested from L1 and L2 DT40 clones which express APOBEC1-Dam and PKA-Dam fusion protein, respectively. (D) This graph illustrates relative methylation of the GC substrates in the endogenous Igλ locus with (L1) and without (L2) Region A. Methylation level of the GC substrates is normalized by that of EF-1α locus. Methylation frequency of after 12 hours AID-Dam induction in L1 clone is set at 1.0. The graph represents the average of three experiments and the average value is indicated below the time points. Error bars represent standard deviations. (E) This histogram compares methylation frequency of the GC substrate by AID-Dam, APOBEC1-Dam and PKA-Dam fusion proteins. The extent of methylation is measured from the genomic DNA harvested after 12 hours induction of each fusion proteins. Methylation level of the gene conversion substrate is normalized by that of EF-1α locus. Methylation level induced by AID-Dam fusion protein is set at 1.0. Abbreviations: E, EcoRI; X, XbaI; M, MboI; Bg, BglII; B, BamHI; P, PvuII; APO, APOBEC1; PKA, Protein kinase A catalytic subunit.

5.3. DISCUSSION

The experiments presented in this chapter suggest that the function of Region A does not lie in creating an open chromatin structure. Instead, Region A appears to be involved in the preferential targeting of AID to the Igλ locus. It has been reported that GC in DT40 cells is stimulated by treatment with the histone deacetylase inhibitor trichostatin A (TSA), which is generally thought to induce permissive chromatin structures (Seo et al., 2005; Lin et al., 2008). However, the effect of TSA is not restricted to the Ig locus, and may therefore affect other factors involved in gene conversion. Given this consideration, it is hard to conclude that permissive chromatin structure *per se* increases gene conversion frequency. Chromatin accessibility of pseudo gene fragments has been reported to affect gene conversion frequency in DT40 cells. Artificial targeting of HP1, which induces heterochromatin silencing, to the pseudo gene region has been reported to cause a dramatic shift in mutation spectrum of Ig gene: reducing templated mutations and promoting point mutations, suggesting that permissive chromatin structure of donor region may play a role in gene conversion activity (Cummings et al., 2007). Consistent with this, tethering VP16, which drives open chromatin structure, to the pseudo gene region increases gene conversion activity in DT40 cells (Cummings et al., 2008). Although these observations suggest that local chromatin accessibility could influence gene conversion, our results would suggest that Region A may not function in this aspect of regulation, but may play a more direct role in recruiting AID to the locus.

It is unclear how Region A facilitates AID recruitment. Since AID has been reported to associate with RNA polymerase II (Nambu et al., 2003), one plausible model is that Region A might play a role in facilitating the interaction between AID and the elongation complex, which in turn deposit AID onto the target template. Region A

could accomplish this function by serving as the binding site for trans-acting factors that associate with AID or it could simply stimulate the levels of transcription elongation through the template.

APPENDUM: Future Directions

A.1. To determine the role of Region A in providing suitable substrate for AID

We showed that Region A is important for AID mediated gene conversion, but has inhibitory effects on I-SceI mediated gene conversion. This result may provide a clue to the mechanism of Region A function. With respect to their substrate requirement, a key distinction between AID and I-SceI is that I-SceI endonuclease cleaves double strand DNA (dsDNA), while AID acts on single strand DNA (ssDNA). It is possible that in the presence of Region A, the GC substrate is present more frequently in an unwound state, thus facilitating AID function, but interfering with I-SceI cleavage. To address this hypothesis, a M13 gene II protein based gene conversion system was established in the DT40 cell line, which expresses the gene II protein in the tetracycline inducible manner (Fig. A.1*B*). The gene II protein is a key initiating factor for plus strand replication in the filamentous bacteriophage, M13 (Horiuchi, 1997). The gene II protein produces a nick at the origin of replication in a site-specific manner. Similar to AID, M13 gene II protein only cleaves single stranded target DNA or negatively supercoiled double stranded DNAs (Horiuchi, 1997). Therefore, M13 gene II protein could serve as a probe for single stranded DNA in our gene conversion substrate. For the substrate of gene II protein, we used the same mutated *puro* reporter system as described above except that we replaced the I-SceI recognition site with the gene II protein recognition sequence (Fig. A.1*A*). If the gene II protein recognizes the sequence and produces a nick, the strand invasion from the nick to the wild type *puro* fragment may induce gene conversion to revert mutated *puro* reporter to a functional *puro* gene. The gene II protein activity is reflected by the gene conversion frequency.

Such gene conversion system has been reported to work in Yeast (Strathern et al., 1991). Using this system, we plan to compare the level of gene conversion mediated by M13 gene II protein in L1 and L2 cells to see if Region A affects the nicking activity of M13 gene II protein, and this result would reflect the levels of strand unwinding at the GC reporter.

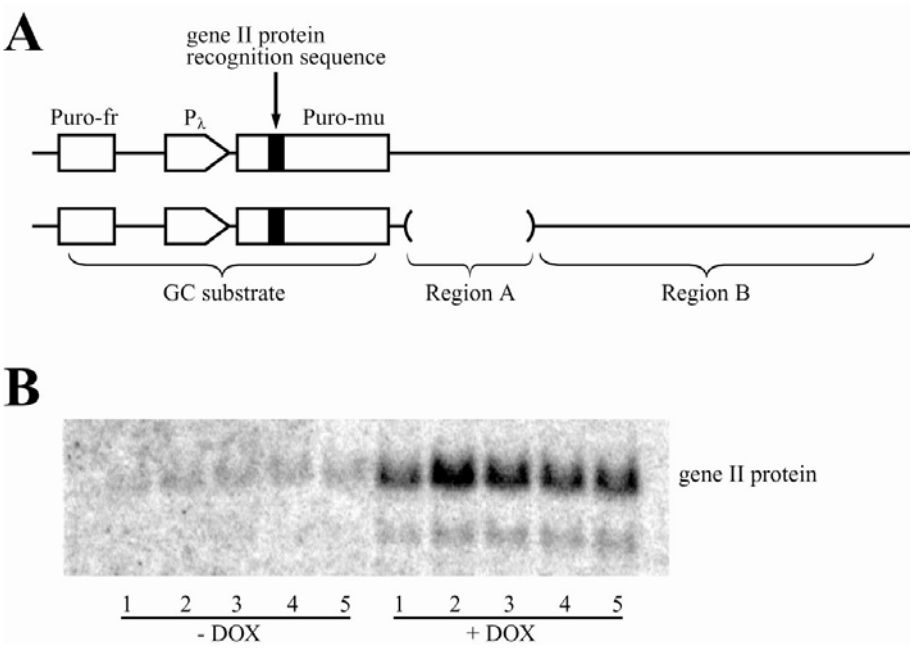

Figure A.1. Gene II protein mediated gene conversion assay

(A) This diagram illustrates gene II mediated gene conversion substrates (L1 and L2) in the Igλ locus. (B) The autoradiograph is Northern blot analysis of gene II protein expression in different DT40 cells before and after induction with doxicycline (Dox). The blot was hybridized with gene II protein cDNA.

A.2. TO DETERMINE THE FUNCTIONAL ROLES OF REGION A BY ENFORCED RECRUITMENT OF AID TO GC SUBSTRATE

Based on the methylation experiments with the AID-Dam fusion, we suggest that Region A facilitates the recruitment of AID to the Igλ locus. To further substantiate the view, we plan to test whether the function of Region A could be replaced with means that artificially target AID to the locus. To test this hypothesis, we introduced an inducible AID-TetR fusion protein construct into the DT40 cell line harboring the GC reporter and TetO without Region A (Fig. A.2*A*). AID-TetR expression depends on cre-mediated recombination upon tamoxifen treatment as described above (Fig. A.2*B*). In this system, the AID-TetR is expected to be recruited to TetO in the GC substrate without Region A. If the enforced recruitment of AID to the GC substrate induces GC activity, this means that recruitment of AID is the limiting step of targeting gene conversion to the GC substrate, and further supports the view that the functions of Region A is to recruit AID to target DNA.

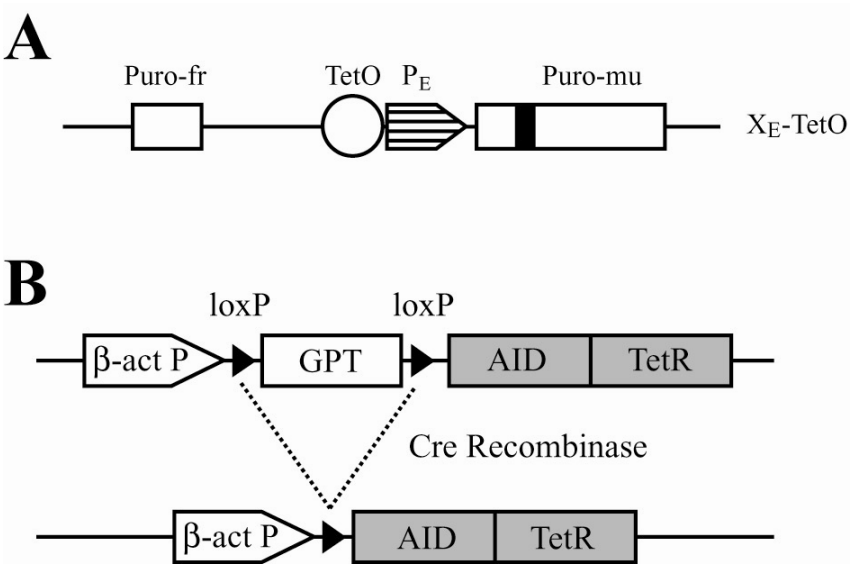

Figure A.2. Enforced recruitment of AID to ectopic gene conversion substrate

(A) This diagram illustrates ectopic gene conversion substrate harboring Tet operator (TetO) upstream of EF-1α promoter. (B) Schematic illustration of inducible AID-Tet repressor (TetR) fusion protein expression.

Bibliography

Arakawa H and Buerstedde JM (2009) Activation-induced cytidine deaminase-mediated hypermutation in the DT40 cell line *Philos Trans R Soc Lond B Biol Sci.* (364) 639-44

Arakawa H and Buerstedde JM (2004) Immunoglobulin gene conversion: insights from bursal B cells and the DT40 cell line *Dev Dyn.* (229) 458-464

Arakawa H, Saribasak H and Buerstedde JM (2004) Activation-induced cytidine deaminase initiates immunoglobulin gene conversion and hypermutation by a common intermediate *PLoS Biol*. (2) 967-974

Arakawa H, Hauschild J, and Buerstedde JM (2002) Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion *Science* (295) 1301-1306

Arakawa H, Lodygin D, and Buerstedde JM (2001) Mutant loxP vectors for selectable marker recycle and conditional knock-outs *BMC Biotechnol.* 1: 7

Bachl J, Carlson C, Gray-Schopfer V, Dessing M and Olsson C (2001) Increased transcription levels induce higher mutation rates in a hypermutating cell line *J. Immunol.* (166) 5051-5057

Bardwell PD et al., (2004) Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice *Nature Immunol.* (5) 224–229

Barreto VM, Ramiro AR, and Nussenzweig MC (2005) Activation-induced deaminase: controversies and open questions *Trends Immunol.* (26) 90-6

Barreto V *et al.*, (2003) C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion *Mol Cell.* (12) 501-508

Beinke S and Ley SC (2004) Functions of NF-kappaB1 and NF-kappaB2 in immune cell biology *Biochem J.* (382) 393-409

Betz AG, Milstein C, González-Fernández A, Pannell R, Larson T and Neuberger MS (1994) Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region *Cell* (77) 239-248

Bezzubova O, Silbergleit A, Yamaguchi-Iwai Y, Takeda S and Buerstedde JM (1997) Reduced X-ray resistance and homologous recombination frequencies in a RAD54$^{-/-}$ mutant of the chicken DT40 cell line *Cell* (89) 185–193

Blagodatski A, Batrak V, Schmidl S, Schoetz U, Caldwell RB, Arakawa H and Buerstedde JM (2009) A cis-acting diversification activator both necessary and sufficient for AID-mediated hypermutation *PLoS Genet.* 5, e1000332

Buerstedde JM, Reynaud CA, Humphries EH, Olson W, Ewert DL and Weill JC (1990) Light chain gene conversion continues at high rate in an ALV-induced cell line *EMBO J.* (9) 921-927

Bulfone-Paus S, Reiners-Schramm L and Lauster R (1995) The chicken immunoglobulin lambda light chain gene is transcriptionally controlled by a modularly organized enhancer and an octamer-dependent silencer *Nucleic Acids Res.* (23) 1997-2005

Buryanov Y and Shevchuk T (2005) The use of prokaryotid DNA methyltransferases as experimental and analytical tools in modern biology *Analytical Biochemistry* (338) 1-11

Chaudhuri J and Alt FW (2004) Class-switch recombination: interplay of transcription, DNA deamination and DNA repair *Nat Rev Immunol.* (4) 541-52

Combriato G and Klobeck H -G (2001) Regulation of human Ig lambda light chain gene expression by NF-kappa B *J Immunol.* (168) 1259-1266

Conlon TM and Meyer KB (2006) The chicken Ig light chain 3'-enhancer is essential for gene expression and regulates gene conversion via the transcription factor E2A *Eur. J. Immunol.* (36) 139-148

Corcoran LM, Karvelas M, Nossal GJ, Ye ZS, Jacks T and Baltimore D (1993) Oct-2, although not required for early B-cell development, is critical for later B-cell maturation and for postnatal survival *Genes Dev.* (7) 570-582

Cummings WJ, Bednarski DW and Maizels N (2008) Genetic variation stimulated by epigenetic modification., *PLoS One* (12):e4075

Cummings WJ, Yabuki M, Ordinario EC, Bednarski DW, Quay S, and Maizels N (2007) Chromatin structure regulates gene conversion, *PLoS Biol* (10):e246

Delker RK, Fugmann SD and Papavasiliou FN (2009) A coming-of-age story: activation-induced cytidine deaminase turns 10, *Nat Immunol.* (10) 1147-53

Di Noia J and Neuberger MS (2002) Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase *Nature* (419) 43-8

Di Noia J and Neuberger MS (2007) Molecular mechanisms of antibody somatic hypermutation *Annu. Rev. Biochem.* (76) 1-22

Di Noia, J and Neuberger MS (2004) Immunoglobulin gene conversion in chicken DT40 cells largely proceeds through an abasic site intermediate generated by excision of the uracil produced by AID-mediated deoxycytidine deamination *Eur. J. Immunol.* (34) 504-508

Di Noia JM and Neuberger MS (2007) Molecular mechanisms of antibody somatic hypermutation *Annu. Rev. Biochem.* (76) 1-22

Di Noia JM, Williams GT, Chan DT, Buerstedde JM, Baldwin GS and Neuberger MS (2007) Dependence of antibody gene diversification on uracil excision *J. Exp. Med.* (204) 3209-3219

Faili A et al., (2004) DNA polymerase η is involved in hypermutation occurring during immunoglobulin class switch recombination *J. Exp. Med.* (199) 265–270

Frey S et al., (1998) Mismatch repair deficiency interferes with the accumulation of mutations in chronically stimulated B cells and not with the hypermutation process. *Immunity* (9) 127–134

Fukita Y, Jacobs H and Rajewsky K (1998) Somatic hypermutation in the heavy chain locus correlates with transcription *Immunity* (9) 105-114

Gorman JR *et al.,* (1996) The Igκ enhancer influences the ratio of Igκ versus Igλ B lymphocytes *Immunity* (5) 241-252

Gordon MS, Kanegai CM, Doerr JR and Wall R (2003) Somatic hypermutation of the B cell receptor genes B29 (Ig beta, CD79b) and mb1 (Ig alpha, CD79a) *Proc. Natl Acad. Sci. USA* (100) 4126–4131

Gottschling DE (1992) Telomere-proximal DNA in Saccharomyces cerevisiae is refractory to methyltransferase activity in vivo *Proc Natl Acad Sci U S A.* (89) 4062-5

Hagman J, Rudin CM, Haasch D, Chaplin D and Storb U (1990) A novel enhancer in the immunoglobulin lambda locus is duplicated and functionally independent of NF-kappa B *Genes Dev.* (4) 978-992

Harris RS, Sale JE, Petersen-Mahrt SK and Neuberger MS (2002) AID is essential for immunoglobulin V gene conversion in a cultured B cell line *Curr. Biol.* (12) 435-438

Horiuchi K (1997) Initiation mechanisms in replication of filamentous phage DNA *Genes to Cells* (2) 425-432

Horton JR, Liebert K, Bekes M, Jeltsch A and Cheng X (2006) Structure and substrate recognition of the Escherichia coli DNA adenine methyltransferase *J Mol Biol.* (358) 559-70

Imai K et al., (2003) Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination *Nature Immunol.* (4) 1023–1028

Inlay MA, Gao HH, Odegard VH, Lin T, Schatz DG and Xu Y (2006) Roles of the immunoglobulin κ light chain intronic and 3' enhancers in Igκ somatic hypermutation *J. Immunol.* (177) 1146-1151

Inlay MA, Tian H, Lin T and Xu Y (2004) Important roles for E protein binding sites within the immunoglobulin kappa chain intronic enhancer in activating Vkappa Jkappa rearrangement *J Exp Med.* (200) 1205-1211

Ito S *et al.,* (2004) Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1 *Proc Natl Acad Sci U S A.* (101) 1975-1980

Janeway CA, Travers P, Walport M and Shlomchik M (2005) Immunobiology - the immune system in health and disease 6 ed., New York: Garland Science Publishing

Jung D et al., (2006) Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus *Annu Rev Immunol,* (24) 541-70

Kim S, Humphries EH, Tjoelker L, Carlson L and Thompson CB (1990) Ongoing diversification of the rearranged immunoglobulin light-chain gene in a bursal lymphoma cell line *Mol. Cell. Biol.* (10) 3224-3231

Kim Y and Tian M (2009) NF-κB family of transcription factor facilitates gene conversion in chicken B cells *Mol. Immunol.* (46) 3283-3291

Kitao H, Kimura M, Yamamoto K, Seo H, Namikoshi K, Agata Y, Ohta K and Takata M (2008) Regulation of histone H4 acetylation by transcription factor E2A in Ig gene conversion *Int Immunol.* (20) 277-84

Klein U *et al.,* (2006) Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination *Nat Immunol.* (7) 773-782

Köntgen F, Grumont RJ, Strasser A, Metcalf D, Li R, Tarlinton D and Gerondakis S (1995) Mice lacking the c-rel proto-oncogene exhibit defects in lymphocyte proliferation, humoral immunity, and interleukin-2 expression *Genes Dev.* (15) 1965-77

Kothapalli N, Norton DD and Fugmann SD (2008) Cutting edge: a cis-acting DNA element targets AID-mediated sequence diversification to the chicken Ig light chain gene locus *J. Immunol.* (180) 2019-2023

Larson ED, Cummings WJ, Bednarski DW and Maizels N (2005) MRE11/RAD50 cleaves DNA in the AID/UNG-dependent pathway of immunoglobulin gene diversification *Mol Cell* (20) 367-75

Lauster R, Reynaud CA, Mårtensson IL, Peter A, Bucchini D, Jami J and Weill JC (1993) Promoter, enhancer and silencer elements regulate rearrangement of an immunoglobulin transgene *EMBO J.* (12) 4615-4623

Lebecque SG and Gearhart PJ (1990) Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter, and 3' boundary is approximately 1 kb from V(D)J gene *J Exp. Med.* (172) 1717-1727

Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH and Schatz DG (2008) Two levels of protection for the B cell genome during somatic hypermutation *Nature* (451) 841-5

Liu M and Schatz DG (2009) Balancing AID and DNA repair during somatic hypermutation *Trends Immunol.* (4) 173-81

Lin W, Hashimoto S, Seo H, Shibata T and Ohta K (2008) Modulation of immunoglobulin gene conversion frequency and distribution by the histone deacetylase HDAC2 in chicken DT40 *Genes to Cells* (13) 255-268

Longerich S, Orelli BJ, Martin RW, Bishop DK and Storb U (2008) Brca1 in immunoglobulin gene conversion and somatic hypermutation *DNA Repair* (7) 253–266

Manis JP, Tian M and Alt FW (2002) Mechanism and control of class-switch recombination *Trends Immunol.* (23) 31-9

Michael N, Shen HM, Longerich S, Kim N, Longacre A and Storb U (2003) The E box motif CAGGTG enhances somatic hypermutation without enhancing transcription *Immunity* (19) 235-242

Morvan CL, Pinaud E, Decourt C, Cuvillier A and Cogné M (2003) The immunoglobulin heavy-chain locus hs3b and hs4 3' enhancers are dispensable for VDJ assembly and somatic hypermutation *Blood* (102) 1421-1427

Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO and Honjo T (1999) Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells *J Biol Chem.* (274) 18470-6

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y and Honjo T (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme *Cell* (102) 553-563

Muschen M et al. (2000) Somatic mutation of the CD95 gene in human B cells as a side-effect of the germinal center reaction *J. Exp. Med.* (192) 1833–1839

Nambu Y *et al.,* (2003) Transcription-coupled events associating with immunoglobulin switch region chromatin *Science* (302) 2137-2140

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

O'Brien RL, Brinster RL and Storb U (1987) Somatic hypermutation of an immunoglobulin transgene in κ transgenic mice *Nature* (326) 405-409

Odegard VH and Schatz DG (2006) Targeting of somatic hypermutation *Nat. Rev. Immunol.* (6) 573-583

Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K and Honjo T (2003) Constitutive expression of AID leads to tumorigenesis *J Exp Med.* (197) 1173-81

Palmer BR and Marinus MG (1994) The Dam and Dcm strains of E.coli-a review *gene* 143 1-12

Pasqualucci L et al. (1998) BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci *Proc. Natl Acad. Sci. USA* (95) 11816–11821

Pasqualucci L et al. (2001) Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas *Nature* (412) 341–346

Pereira SG and Oakley F (2008) Nuclear factor-kappaB1: regulation and function *Int J Biochem Cell Biol* (40) 1425-30

Perlot T, Alt FW, Bassing CH, Suh H and Pinaud E (2005) Elucidation of IgH intronic enhancer functions via germ-line deletion *Proc. Natl. Acad. Sci. U S A.* (102) 14362-14367

Peters A and Storb U (1996) Somatic hypermutation of immunoglobulin genes is linked to transcription initiation *Immunity* (4) 57-65

Phung QH *et al*., (1998) Increased hypermutation at G and C nucleotides in immunoglobulin variable genes from mice deficient in the MSH2 mismatch repair protein *J. Exp. Med.* (187) 1745–1751

Rada C *et al.*, (2002) Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG deficient mice *Curr. Biol.* (12) 1748–1755

Rada C, Ehrenstein MR, Neuberger MS and Milstein C (1998) Hotspot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting *Immunity* (9) 135–141

Rada C, Yélamos J, Dean W and Milstein C (1997) The 5' hypermutation boundary of kappa chains is independent of local and neighbouring sequences and related to the distance from the initiation of transcription *Eur. J. Immunol.* (27) 3115-3120

Ratcliffe MJ (2006) Antibodies, immunoglobulin genes and the bursa of Fabricius in chicken B cell development *Dev Comp Immunol,* (30) 101-18

Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R, Gennery A, Tezcan I, Ersoy F, Kayserili H, Ugazio AG, Brousse N, Muramatsu M, Notarangelo LD, Kinoshita K, Honjo T, Fischer A and Durandy A (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2) *Cell* (102) 565-575

Reynaud CA *et al*., (1987) A hyperconversion mechanism generates the chicken light chain preimmune repertoire *Cell* (48) 379-88

Reynaud CA, Dahan A, Anquez V and Weill JC (1989) Somatic hyperconversion diversifies the single Vh gene of the chicken with a high incidence in the D region *Cell* (59) 171-183

Reynaud CA, Dahan A and Weill JC (1983) Complete sequence of a chicken lambda light chain immunoglobulin derived from the nucleotide sequence of its mRNA *Proc Natl Acad Sci U S A.* (80) 4099-4103

Robbiani DF, Bothmer A, Callen E, Reina-San-Martin B, Dorsett Y, Difilippantonio S, Bolland DJ, Chen HT, Corcoran AE, Nussenzweig A and Nussenzweig MC (2008) AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations *Cell* (135) 1028-38

Ronai D, Iglesias-Ussel MD, Fan M, Shulman MI and Scharff MD (2005) Complex regulation of somatic hypermutation by cis-acting sequences in the endogenous IgH gene in hybridoma cells *Proc. Natl. Acad. Sci. U S A.* (102) 11829-11834

Rooney S, Chaudhuri J and Alt FW (2004) The role of the non-homologous end-joining pathway in lymphocyte development *Immunol Rev.* (200) 115-31

Rouet P, Smih F and Jasin M (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease *Mol Cell Biol*. (14) 8096-8106

Sale JE (2004) Immunoglobulin diversification in DT40: a model for vertebrate DNA damage tolerance. *DNA Repair* (3) 693-702

Sale JE, Calandrini DM, Takata M, Takeda S and Neuberger MS (2001) Ablation of XRCC2/3 transforms immunoglobulin V gene conversion into somatic hypermutation *Nature* (412) 921-926

Saribasak H, Saribasak NN, Ipek FM, Ellwart JW, Arakawa H and Buerstedde JM (2006) Uracil DNA glycosylase disruption blocks Ig gene conversion and induces transition mutations *J. Immunol.* (176) 65-371

Satyaraj E and Storb U (1998) Mef2 proteins, required for muscle differentiation, bind an essential site in the Ig lambda enhancer *J Immunol.* (161) 4795-4802

Schoetz U, Cervelli M, Wang YD, Fiedler P and Buerstedde JM (2006) E2A expression stimulates Ig hypermutation *J. Immunol*. (177) 395-400

Schrader CE *et al.*, (1999) Reduced isotype switching in splenic B cells from mice deficient in mismatch-repair enzymes *J. Exp. Med.* (190) 323–330

Sen R, (2004) NF-kappaB and the immunoglobulin kappa gene enhancer *J Exp Med.* (200) 1099-1102

Seo H, Masuoka M, Murofushi H, Takeda S, Shibata T, *et al*. (2005) Rapid generation of specific antibodies by enhanced homologous recombination *Nat Biotechnol* (23) 731–735

Sha WC, Liou HC, Tuomanen EI and Baltimore D (1995) Targeted disruption of the p50 subunit of NF-kappa B leads to multifocal defects in immune responses *Cell* (80) 321-30

Shen HM, Peters A, Baron B, Zhu X and Storb U (1998) Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes *Science* (280) 1750–1752

Shen HM, Poirier MG, Allen MJ, North J, Lal R, Widom J and Storb U (2009) The activation-induced cytidine deaminase (AID) efficiently targets DNA in nucleosomes but only during transcription *J Exp Med* (206) 1057-71

Shinkura R *et al.* (2004) Separate domains of AID are required for somatic hypermutation and class-switch recombination *Nat Immunol.* (5) 707-712

Shivarov V, Shinkura R, Doi T, Begum NA, Nagaoka H, Okazaki IM, Ito S, Nonaka T, Kinoshita K and Honjo, T (2009) Molecular mechanism for generation of antibody memory *Philos Trans R Soc Lond B Biol Sci.* (364) 569-575

Singh J and Klar AJ (1992) Active gene s in budding Yeast display enhanced in vivo accessibility to foreign DNA methylase: a novel in vivo probe for chromatin structure of Yeast *Genes Dev*. (6) 186-196

Steensel BV and Henikoff S (2000) Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase *Nature Biotechnology* (18) 424-328

Strathern JN, Weinstock KG, Higgins DR and McGill CB (1991) A novel recombinator in Yeast based on gene II protein from bacterophage f1 *Genetics* (127) 61-73

Swanson BJ, Jäck HM and Lyons GE (1998) Characterization of myocyte enhancer factor 2 (Mef2) expression in B and T cells: Mef2C is a B cell-restricted transcription factor in lymphocytes *Mol Immunol* (35) 445-58

Ta VT *et al.* (2003) AID mutant analyses indicate requirement for class-switch-specific cofactors *Nat Immunol.* (4) 843-848

Takizawa M, Tolarová H, Li Z, Dubois W, Lim S, Callen E, Franco S, Mosaico M, Feigenbaum L, Alt FW, Nussenzweig A, Potter M and Casellas R (2008) AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development *J Exp Med.* (205) 1949-57

Tumas-Brundage K and Manser T (1997) The transcriptional promoter regulates hypermutation of the antibody heavy chain locus *J. Exp. Med.* (185) 239-250

van der Stoep N, Gorman JR and Alt FW (1998) Reevaluation of 3'Eκ function in stage- and lineage-specific rearrangement and somatic hypermutation *Immunity* (8) 743-750

van der Stoep N, Gorman JR and Alt FW (1998) Reevaluation of 3'Eκ function in stage- and lineage-specific rearrangement and somatic hypermutation *Immunity* (8) 743-750

Venetianer A, Schiller DL, Magin T and Franke WW (1983) Cessation of cytokeratin expression in a rat hepatoma cell line lacking differentiated functions *Nature* (305) 730-3

Wang VE, Tantin D, Chen J and Sharp PA (2004) B cell development and immunoglobulin transcription in Oct-1-deficient mice *Proc Natl Acad Sci U S A.* (101) 2005-2010

Wiesendanger M, Kneitz B, Edelmann W and Scharff MD (2000) Somatic hypermutation in MutS homologue (MSH)3-, MSH6-, and MSH3/MSH6-deficient mice reveals a role for the MSH2–MSH6 heterodimer in modulating the base substitution pattern *J. Exp. Med.* (191) 579–584

Wright JH, Gottschling DE and Zakian VA (1992) Saccharomyces telomeres assume a non-nucleosomal chromatin structure *Genes Dev*. (6) 197-210

Xiang Y and Garrard WT (2008) The Downstream Transcriptional Enhancer, Ed, positively regulates mouse Ig kappa gene expression and somatic hypermutation *J. Immunol.* (180) 6725-6732

Yabuki M, Fujii MM and Maizels N (2005) The MRE11-RAD50-NBS1 complex accelerates somatic hypermutation and gene conversion of immunoglobulin variable regions *Nat Immunol.* (6) 730-6

Yamamoto K *et al.* (2005) Fanconi anemia protein FANCD2 promotes immunoglobulin gene conversion and DNA repair through a mechanism related to homologous recombination *Mol. Cell. Biol.* (25) 34–43

Yang SY, Fugmann SD and Schatz DG (2006) Control of gene conversion and somatic hypermutation by immunoglobulin promoter and enhancer sequences *J. Exp. Med.* (203) 2919-2928

Yelamos J, Klix N, Goyenechea B, Lozano F, Chui YL, González Fernández A, Pannell R, Neuberger MS and Milstein C (1995) Targeting of non-Ig sequences in place of the V segment by somatic hypermutation *Nature* (376) 225-229

Zeng X *et al.* (2001) DNA polymerase η is an A•T mutator in somatic hypermutation of immunoglobulin variable genes *Nature Immunol.* (2) 537–541

Vita

Yonghwan Kim was born in Seoul, Korea on April 30, 1975, the son of Hongsik Kim and Kangsoon Lee. After completing his work at Hansung high school, in 1994 he entered Seoul National University in Seoul, Korea. He received the degrees of Bachelor of Science in 1998, and Master of Science in 2000 from Seoul National University. In 2000, he joined the Republic of Korea Army and served for twenty six months. In August, 2004, he entered the graduate school of the University of Texas at Austin as a student in the section of molecular genetics and microbiology.

Permanent address: Hyundai APT 107-704, Cheolsan Dong, Gwangmyeong Si
Gyeongki Do, 423-721 (zip)
South Korea

This dissertation was typed by the author.