12-14-2010

The C Terminus of Activation Induced Cytidine Deaminase (AID) Recruits Proteins Important for Class Switch Recombination to the IG Locus: A Dissertation

Sanjay Ranjit
University of Massachusetts Medical School Worcester, Sanjay.Ranjit@umassmed.edu

THE C TERMINUS OF ACTIVATION INDUCED CYTIDINE DEAMINASE (AID) RECRUITS PROTEINS IMPORTANT FOR CLASS SWITCH RECOMBINATION TO THE IG LOCUS

A Dissertation Presented

By

Sanjay Ranjit

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences,

Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 14, 2010

MOLECULAR GENETICS AND MICROBIOLOGY

THE C TERMINUS OF ACTIVATION INDUCED CYTIDINE DEAMINASE (AID) RECRUITS PROTEINS IMPORTANT FOR CLASS SWITCH RECOMBINATION TO THE IG LOCUS

A Dissertation Presented By

Sanjay Ranjit

The signature of the Dissertation Defense Committee signifies completion and approval as to style and content of the Dissertation

Janet Stavnezer, PhD., Thesis Advisor

Nicholas Rhind, PhD., Member of Committee

Michael Volkert, PhD., Member of Committee

Robert Woodland, PhD., Member of Committee

Erik Selsing, PhD., Member of Committee

The signature of the Chair of the committee signifies that the written dissertation meets the requirements of the Dissertation Committee

Rachel Gerstein, PhD., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school.

Anthony Carruthers, PhD., Member of Committee
Dean of the Graduate School of Biomedical Sciences

Molecular Genetics and Microbiology
Decemeber 14, 2010

Acknowledgements

I would like to thank my mentor, Janet Stavnezer for her outstanding support and guidance. I would like to thank current and previous members of Stavnezer lab for all the help and advice; Carol Schrader for reading my manuscript and helping with preparation for talks, Anna Ucher for helping to prepare cell cultures/retroviruses, Erin Linehan for sequence analysis of LM-PCR products. I also would like to thank my committee members Rachel Gerstein, Nick Rhind, Mike Volkert and Robert Woodland for their constructive advice,and Erik Selsing for joining my committee and coming to UMASS for my defense. Finally, I would like to thank all my family and friends, who supported me through the years.

Abstract

Activation-induced cytidine deaminase (AID) is a key protein required for both class switch recombination (CSR) and somatic hypermutation (SHM) of antibody genes. AID is induced in B cells during an immune response. Lack of AID or mutant form of AID causes immunodeficiency; e.g., various mutations in the C terminus of AID causes hyper IgM (HIGM2) syndrome in humans. The C terminal 10 amino acids of AID are required for CSR but not for SHM. During both CSR and SHM, AID deaminates dCs within Ig genes, converting them to dUs, which are then either replicated over, creating mutations, or excised by uracil DNA glycosylase (UNG), leading to DNA breaks in Ig switch regions. Also, the mismatch repair (MMR) heterodimer Msh2-Msh6 recognizes U:G mismatches resulting from AID activity and initiates MMR, which leads to increased switch region double strand breaks (DSBs). DSBs are essential intermediates of CSR; lack of UNG or MMR results in a reduction of DSBs and CSR. The DSBs created in the Sμ and one of the downstream S-regions during CSR are recombined by non-homologous end joining (NHEJ) to complete CSR. Available data suggest that AID is required not only for the deamination step of CSR, but also for one or more of the steps of CSR that are downstream of deamination step. This study investigates the role of C terminus of AID in CSR steps downstream of deamination.

Using retroviral transduction into mouse splenic B cells, I show that AID binds cooperatively with UNG and Msh2-Msh6 to the Ig Sμ region, and this depends on the AID C terminus. I also show that the function of MMR during CSR depends on the AID C terminus. Surprisingly, the C terminus of AID is not required for Sμ or Sγ3 DSBs, suggesting its role in CSR occurs during repair and/or recombination of DSBs.

Table of Contents

Approval page……ii

Acknowledgement……iii

Abstract……iv

List of Figures……viii

List of Frequently used Abbreviations……ix

CHAPTER I: Introduction……1

Double-strand Break formation during CSR……4

Somatic Hypermutation……10

AID protein and deamination activity……11

Regulation of AID……15

Transcriptional regulation……15

Post-transcriptional/post-translational regulation……16

Protein-Protein (AID) interactions……17

C-terminus of AID……19

Hypothesis……20

CHAPTER II: Materials and methods.....24

CHAPTER III: Results……37

Cell cycle progression and cell proliferation……37

Co-dependent binding of AID and UNG with Sμ region……...42

Co-dependent binding of AID and Msh2-Msh6 with Sμ……..…46

Role of C-terminus of AID for the function of MMR during CSR……50

Dominant negative effect of ΔAID during CSR……..……54

AID and UNG do not interact *in vitro*……...59

Role of C-terminus of AID for DSB formation during CSR……..62

CHAPTER IV: Discussion……..……66

Future directions……76

Bibliography……..……78

List of Figures

Figure 1 CSR model……..…..2

Figure 2 Role of AID during CSR and SHM……....6

Figure 3 MMR model for DSB formation during CSR……..…..8

Figure 4 Schematic of AID protein……...13

Figure 5 Experimental approach……..22

Figure 6 Investing getting the role of ΔAID in cell cycle and cell proliferation of *aid-/-* splenic B cells……...40

Figure 7 ChIP with anti-ER and anti-UNG antibodies……..44

Figure 8 ChIP with anti-Msh2 and anti-Msh6 antibodies……....48

Figure 9 CSR assay demonstrating importance of C-terminus of AID for the function of MMR proteins……52

Figure 10 Dominant negative effect of ΔAID during CSR……57

Figure 11 AID and UNG do not interact *in vitro*……60

Figure 12 DSB phenotype of *aid-/-* cells expressing ΔAID……..…64

Figure 13 Model for AID complex formation during CSR……..…74

List of Frequently used Abbreviations

AID	Activation induced cytidine deaminase
ΔAID	last 10 C-terminal amino acid deleted AID mutant
AID-ER	AID and estrogen receptor fusion protein (fused at the C-terminus of AID)
RV-AID-ER	Retroviral construct containing AID-ER expression cassette
CSR	Class Switch Recombination
SHM	Somatic Hypermutation
UNG	Uracil DNA Glycosylase
MMR	Mismatch Repair Protein
DSBs	Double Strand Breaks
APE	Apurinic-Apyrimydinic Endonuclease
MMR	Mismatch Repair
Msh2	MutS homolog2
Msh6	MutS homolog6
MLH1	MutL homolog1
NHEJ	Non-homologous End Joining
PKA	cAMP dependent Protein Kinase A
RPA	Replication Protein A
HIGM2	Hyper-IgM2 syndrome

DNA-PKcs	DNA-dependent protein kinase, catalytic subunit

Chapter I

INTRODUCTION

When naïve B cells are activated via infection or immunization, they switch immunoglobulin (Ig) heavy chain isotypes, from initial expression of IgM and IgD to expression of IgG, IgE, or IgA, which alters the effector functions of the antibodies. Isotype switching occurs by class switch recombination (CSR), a unique intrachromosomal (mostly) DNA recombination, which involves introduction of double-strand breaks (DSBs) into special DNA regions, termed switch (S) regions, which range from 1 to 12 kb in length, and are located directly upstream of each constant region (C) gene of the *Igh* locus (1). During B cell activation, single nucleotide mutations are also introduced into antibody variable regions in a process called somatic hypermutation (SHM).

Activation-induced cytidine deaminase (AID) initiates CSR and SHM. AID deaminates dCs to dUs, preferentially in WRC motifs (where W = A or T; R = G or A) (2-5). These motifs are very abundant in S region DNA. During CSR, the dUs created by AID activity at S regions are converted into DSBs in the donor (Sμ) and acceptor Sx regions, which are then recombined typically by non-homologous end joining (NHEJ), and the intervening sequence is removed as a circular product (Fig 1).

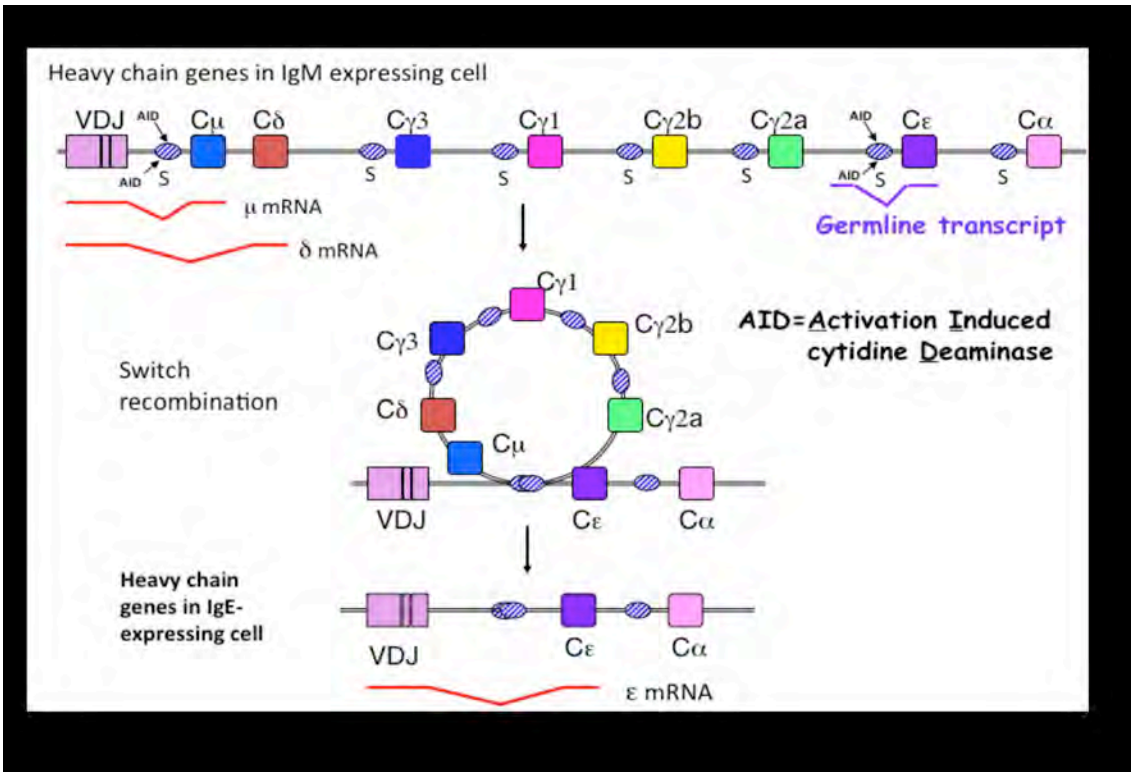
Heavy chain genes in IgM expressing cell
VDJ
AID
Cμ
Cδ
Cγ3
Cγ1
Cγ2b
Cγ2a
Cε
Cα
S
μ mRNA
δ mRNA
Germline transcript
AID=Activation Induced cytidine Deaminase
Switch recombination
Heavy chain genes in IgE-expressing cell
ε mRNA

Figure 1

Figure 1 CSR model. *Igh* locus showing constant region genes for IgM, IgD, IgG3, IgG1, IgG2b, IgG2a, IgE and IgA after a VDJ recombination. Each C-region is preceded by an S-region, where transcription takes place to generate AID substrates, which also generates germline transcripts: Sμ though Cμ, Sμ though Cδ, and Sε through Cε (in this particular case). AID induces DSBs in the S –regions; which S-region depends on the mode of activation. The DSB on Sμ and one of the downstream Sx region are then joined by NHEJ. The intervening sequence is removed as a circular product. On this figure, the end result is switching to IgE (1).

Double-strand break formation during CSR

The dUs generated during CSR and SHM on S regions and V genes, respectively, are removed by uracil-*N*-glycosylase2 (UNG). Deficiency in UNG results in a dramatic reduction in CSR, chromosomal translocations between the *IgH* and *c-myc* loci, and causes a dramatic increase in G>A and C>T transition mutations during SHM as DNA is replicated without UNG being removed (6-9). Apurinic-apyrimidinic endonuclease 1 and 2 (APE1/2) have been shown to be important for DSB formation during CSR (10), presumably due to their ability to nick the abasic sites generated by UNG activity during CSR. If the nicked sites on opposite strands are sufficiently near, they can spontaneously form DSBs that are substrates for NHEJ (Fig 2).

On the other hand, AID generated G:U mismatches are also substrates for mismatch repair proteins (MMR). If the single stranded nicks on opposite DNA strands, generated by AID-UNG-APE activities, are too far apart, MMR proteins can process the single-stranded nicks into DSBs during CSR (10-12). The (MutS homologues) Msh2-Msh6 heterodimer can bind to G:U mismatches, then recruit Mlh1-Pms2 heterodimers (13, 14). Binding of these MMR proteins to the mismatch promotes loading of Exonuclease 1 (Exo1) (15) at the SSBs that are generated by APE1/2 activity. Exo1 can continue excision past the U:G mismatch until reaching a SSB on the other DNA strand, thus creating a DSB with staggered ends. A DNA polymerase can then fill in the staggered end to

generate a blunt or nearly blunt-ended DSB (11) (Fig 3).. If the nearest SSB is located 3’ to the U:G mismatch, the nicking activity of Pms2 is required to create a SSB 5’ to the mismatch in order for Exo1 to be able to excise it (16, 17).

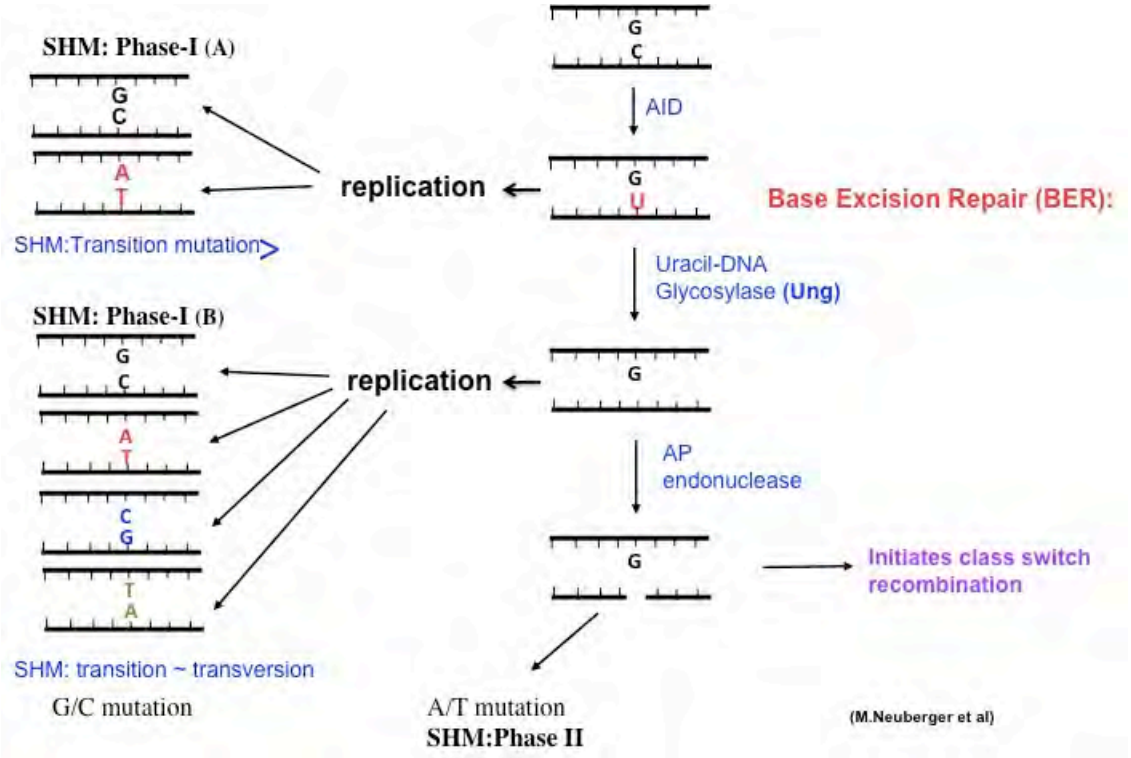
SHM: Phase-I (A)
G
C
A
T
SHM:Transition mutation
replication
G
C
AID
G
U
Base Excision Repair (BER):
Uracil-DNA
Glycosylase (Ung)
SHM: Phase-I (B)
G
C
A
T
C
G
T
A
SHM: transition ~ transversion
replication
G
AP
endonuclease
G
Initiates class switch
recombination
G/C mutation
A/T mutation
SHM:Phase II
(M.Neuberger et al)

Figure 2

Figure 2 AID is required for both CSR and SHM. AID deaminates dCs to dUs during CSR and SHM. During SHM, the dUs can be replicated to produce G:C to A:T transition mutations, which is termed as phase-I(A) of SHM; or the dUs can be removed by UNG. The abasic site can be replicated to produce transition/transversion mutations, which is termed as phase-I(B) of SHM (18). The abasic site can also be nicked by AP endonuclease, an error prone polymerase such as polymerase η can introduce a patch with mutations at A:T base pairs via strand-displacement reaction (19), this is termed as phase-II of SHM. During CSR, some of these nicks are eventually converted into DSBs.

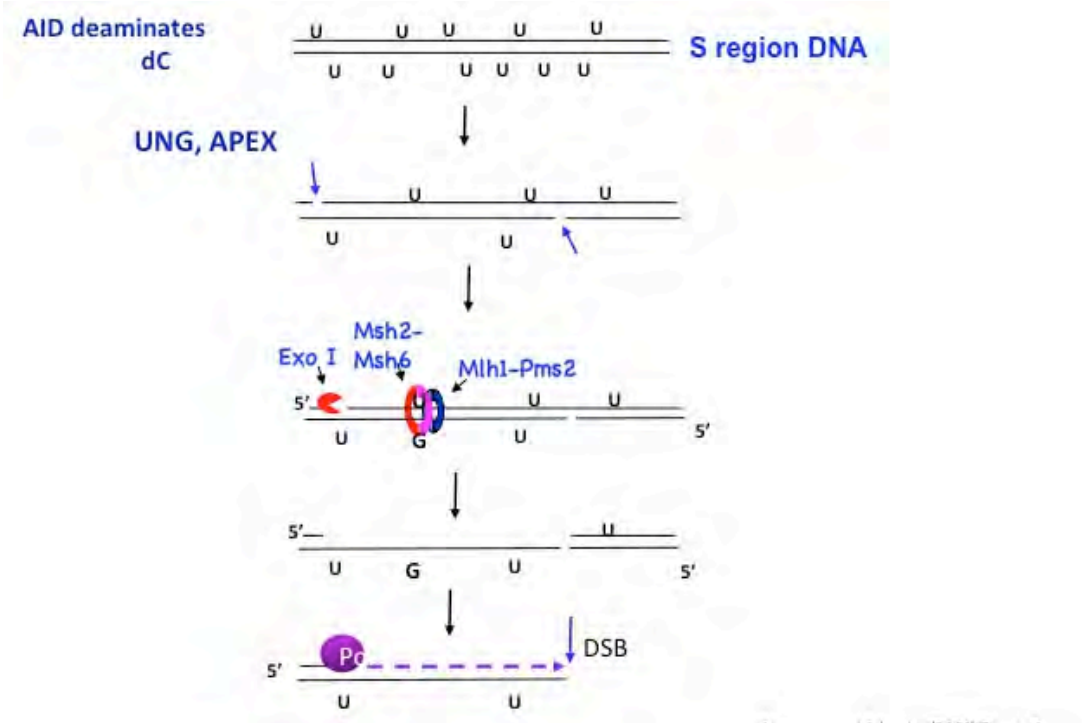
AID deaminates
dC
S region DNA
UNG, APEX
Msh2-
Msh6
Exo I
Mlh1-Pms2
5'
5'
G
DSB

Figure 3

Figure 3 MMR model for DSB formation during CSR. The Msh2-Msh6 heterodimer can bind to the U:G mismatch and recruit Exo I to a nicked site located 5' side of the U:G mismatch. The Exo I can remove a patch of DNA until it reached the nicked site on the other strand. A DNA polymerase then fills in the staggered end of the DNA to generate a DSB (1).

SHM

Unlike CSR, SHM induces single amino acid mutations, most frequently at the antigen recognition sites of antibodies (complementarity determining regions, CDRs), which can enhance the ability of an antibody to bind antigens (20-22). During SHM, error prone polymerases can replicate DNA through the abasic sites created by UNG activity to generate transition and transversion mutations (8, 18). At the same time, the DNA with dUs can be replicated to generate G>A and C>T transition mutations. In *ung-/-* mice, G:C mutations are severely skewed towards G>A and C>T transition mutations, while the A:T mutations are similar in both *ung-/-* and WT mice (6, 7, 23). Hence, the AID and UNG activities at G:C base pairs generate the first phase of SHM as postulated by Petersen-Mahrt, S *et al* (18) (Fig 2).

In the second phase of SHM, G:U mismatches induced by AID can be recognized by the mismatch repair proteins, Msh2 and Msh6 heterodimer (21). As described above, this will in turn recruit MutL homologues (Mlh1 and Pms2 heterodimer), and Exo1 to the mismatch. Mice deficient in Msh2 (24-26) or Msh6 (27-29) show very few A:T mutations in V regions in mice. Interestingly, it has been also shown that *msh6-/-* mice have an altered targeting of AID, where the AID induced mutations in V regions are more focused to the AID target WRC/GYW hotspots (28, 30). Hence, actions of AID and MMR are very important to generate mutations at both G:C and A:T base pairs during the

second phase of SHM, and MMR proteins have a role in not only steps downstream of AID activity, but also in targeting of AID itself.

AID protein and deamination activity

Mouse AID is a 198 amino acid, 24 kD protein; mouse AID has 92% homology to human AID (31). AID has a putative nuclear localization signal at the N-terminus between amino acids 9 and 26 (32), a SHM domain between amino acids 13 and 23 (33), a catalytic (deaminase) domain between amino acids 56 and 94, and a domain required for CSR, but not SHM, between amino acids 182 and 198 (34-37). It also has a nuclear export signal (NES) at the C-terminus between amino acids 190 and 198 (32, 38, 39) (Fig 4).

AID has been shown to preferentially deaminate dCs to dUs in WRC motifs on single stranded DNA *in vitro (2-5, 40, 41)*. Transcription through S regions is a prerequisite for CSR, apparently due to the ability of transcription to generate single stranded (ss) DNA (42, 43), the substrate for AID (3, 4, 44, 45). Also, AID interacts with RNA polymerase II (RNPII), suggesting that AID targeting (46) to the *Ig* locus is linked to the transcription complex during CSR and SHM (47). Transcription also induces histone modifications that might increase accessibility of the DNA to AID (48, 49).

According to *in vitro* experiments, AID has a higher activity when it is in a higher order oligomeric form (50). The dimerization domain of AID is unknown, but C-terminus truncated AID (last 10 amino acids) (ΔAID) can still form an oligomer. Also, human patients heterozygous for Δ AID present with HIGM2 syndrome, indicating that Δ AID might have a dominant negative effect (36, 51). It is also important to note that these patients have normal SHM levels (36, 51).

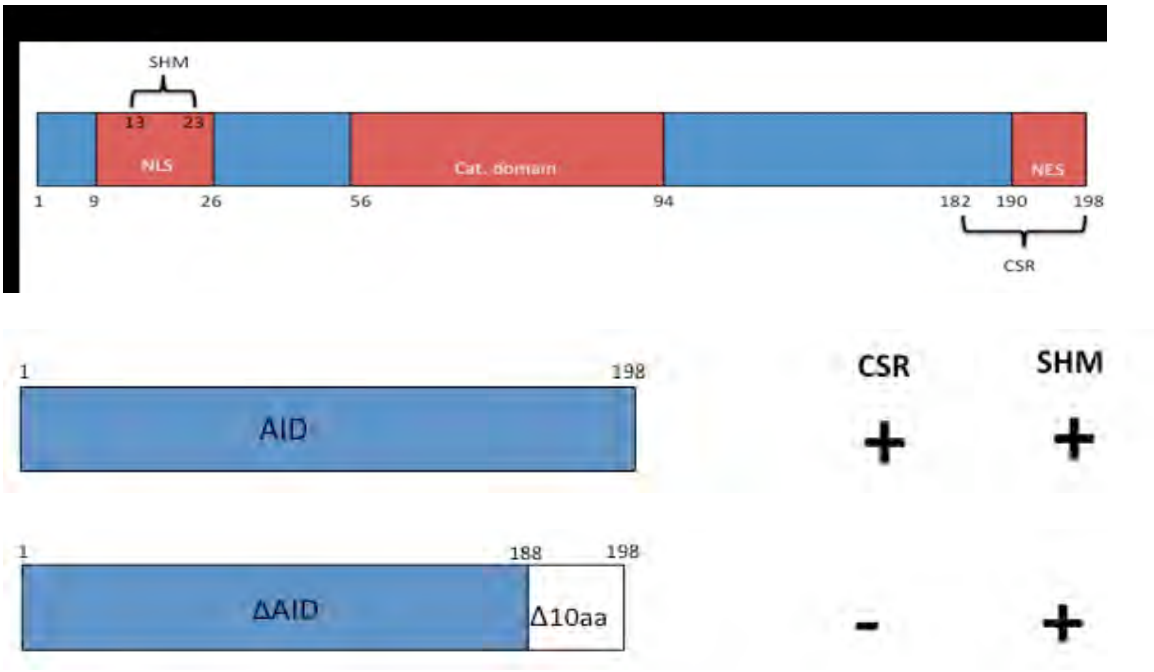
SHM
13
23
NLS
Cat. domain
NES
1
9
26
56
94
182
190
198
CSR
1
198
AID
CSR
SHM
+
+
1
188
198
ΔAID
Δ10aa
-
+

Figure 4

Figure 4 Schematic representation of the structure of AID; NLS (nuclear localization signal) domain, SHM (somatic hypermutation) domain, catalytic/deamination domain, CSR (class switch recombination) domain, and NES (nuclear export signal) domain. Last 10 amino acids in the C-terminus of AID are required for CSR but not SHM.

Regulation of AID

Although AID is essential for CSR and SHM, tight regulation of AID activity is imperative as AID activity is potentially highly genotoxic. Recent developments suggest that although AID is expressed predominantly in germinal center B cells, it is expressed in a variety of other tissues at much lower levels as well (52, 53). Mouse models constitutively expressing AID are prone to developing tumors, e.g. lung tumors and T-cell lymphomas (54, 55). However, a transgenic mouse model with AID expression driven by the CD19 promoter does not develop B cell lymphomas (56), suggesting that B cells negatively regulate AID. In addition, AID has been implicated in causing *c-myc* and *Igh* translocations similar to those in lymphomas (9, 57, 58). Although the ability of AID to deaminate dCs only on ssDNA appears to limit its activity to transcribed genes, it might induce DNA lesions in an array of non-*Ig* genes transcribed during CSR.

Transcriptional regulation of AID

AID expression in B cells is induced by the same signals that activate B cells to undergo CSR. *In vitro,* AID is expressed in mouse B cells 48 hours after cytokine activation (42). The AID genetic locus (*aicda*) has four regions that are found to be important for transcriptional regulation, a 1.5 kb promoter with NF-κB, STAT6, HoxC4, Sp1 and Pax5 binding sites, but no TATA box (59-62). Signal transduction pathways such as JAK/STAT and signals downstream of NF-κB have been shown to be involved in AID expression in B cells (61, 63). E47, an E-

protein, a helix-loop-helix transcription factor, can activate the intronic enhancer in the *aicda* gene to induce AID expression in B cells (60). Blimp1, a transcription factor involved in plasma cell differentiation, which occurs after CSR and SHM, inhibits the expression of the transcription factor Pax5 (64), as well as AID (65). Also, interferon regulatory factor (IRF-4), which inhibits Blimp-1 transcription, increases AID expression (66, 67).

Post-transcriptional/post-translational regulation of AID

AID is post-transcriptionally negatively regulated at the mRNA level by the lymphocyte specific micro-RNAs, miR-155 and miR-181b (68-70). Also, post-translational modification of nuclear AID at Ser38 via phosphorylation was shown to be important for both CSR and SHM (71-74). AID protein level in the nucleus of B cells itself is limited, and AID is suggested to have a weak nuclear localization signal (NLS) on its N-terminal domain (32, 38, 39). By contrast, the C-terminus of AID has a robust nuclear export signal (NES) dependent upon chromosome region maintenance-1 (CRM1) protein (39). Hence, AID protein predominantly resides in the cytoplasm (75). Although the C terminus of AID is also required for CSR, nuclear export is not required for CSR, as mouse AID^{F198A} is not exported from nuclei, and yet CSR and SHM are only modestly reduced in cells expressing this mutant (39). Therefore, restrictive concentration of AID protein in the nucleus by weak nuclear localization and CRM1-mediated nuclear export of AID has been suggested to curb AID induced lesions in non-*Ig* loci (39).

The activity of AID seems to be further limited by enhanced ubiquitin-mediated degradation of nuclear AID. The half-life of nuclear AID is shorter, approximately three times less than that of cytoplasmic AID (76). However, since patients heterozygous or homozygous for mutations in the NES of AID have no change in AID-induced mutation frequencies (34, 36, 51), this indicates that lengthening the life span of nuclear AID does not increase mutagenesis. Although ubiquitin-mediated proteasomal degradation of AID is not fast enough to completely shut down AID deamination activity, it might prevent AID protein from carrying out steps downstream of deamination during CSR in *Igh* locus such as deletion recombination (CSR like event) in non-*Ig* loci . Hence, ubiquitin-mediated proteasomal degradation of AID might help to fine-tune AID regulation by reducing the concentration AID protein in the nucleus.

AID is involved in protein-protein interactions

It has been suggested that protein-protein interactions are important for AID to perform CSR, because several AID mutants have defects in CSR, in particular, some with the mutations in NES are unable to retain CSR while they retain SHM (34, 76, 77). Hence, the ability of AID to form a CSR specific complex via protein-protein interaction might be another layer of protection from AID induced genomic modification such as inter-gene recombination to prevent translocations or harmful deletion of genes/parts of genes.

AID has been shown to interact with numerous proteins, but no CSR specific or SHM specific factor that interacts with AID has been published yet, except for 14-3-3 adaptor protein and DNA-dependent protein kinase catalytic subunit (DNA-PKcs). A recent publication reported that the C-terminus of AID interacts with the 14-3-3 adaptor protein complex and that this interaction is important for recruitment of AID to S regions (78). Deficiency in 14-3-3 protein components results in a 50% decrease in CSR. DNA-PKcs has been shown to interact with the C-terminus of AID in a DNA dependent manner (79). Although DNA-PKcs has been suggested to be important for joining AID dependent DSBs in *Igh* locus (80), DNA-PKcs deficient mice have not consistently shown impairment in CSR. AID has also been shown to interact with Spt5, thereby bringing AID to sites of RNPII stalling (81). Also, AID interacts with RNA polymerase II (RNPII), suggesting that AID targeting to the *Igh* locus is linked to the transcription complex during CSR and SHM (47). This raises more questions because AID-RNPII interaction does not provide *Igh* locus specificity for AID; instead it might make all the genes transcribed in activated B cells vulnerable to AID activity. It has also been shown that AID interacts with CTNNBL1, a component of the RNA splicing complex. CTNNBL1 is suggested to be one of the AID interacting proteins that might be involved in targeting of AID to the *Igh* locus (82). However, deletion of this protein does not affect CSR (83). Another protein, MDM2, an oncoprotein, which targets p53 for nuclear export and degradation,

was shown to interact with the C-terminus of AID, but MDM2 is not important for antibody diversification (84).

The regulatory and catalytic subunit of protein kinase A (PKA), which phosphorylates AID at S38, is specifically localized in S region DNA during CSR, and might participate in a PKA-RPA-AID complex (85). Phosphorylation of S38 on AID is not required for S region binding of AID but it is required for CSR; phosphorylation of S38 does promote interaction between AID and Replication Protein A32 (RPA32), which enhances the DNA binding property of AID, and also its deaminase activity (71, 86). Hence, it is suggested that PKA causes localized phosphorylation of AID and nucleates a CSR-specific complex on the S region (85). Such a CSR-specific complex might be important for promoting acceptor/donor S-region DNA synapses and/or facilitating NHEJ, making these processes mediated by AID associated with S-region DNA.

The C-terminus of AID is important for CSR

Some of the NES mutants that are unable to support CSR are still able to perform SHM of Ig V regions (34, 76, 77). For example, mouse B cells expressing ΔAID (last 10 amino acids deleted) have greatly reduced CSR (~10% of WT), but retain normal levels of SHM and AID-induced mutations in the Sμ region (34). Although the mutation frequencies in the Sμ region are unaltered in ΔAID expressing B cells, G>A and C>T transition mutations are increased by two

fold compared to that of WT AID expressing B cells (34), which is similar to the effect of UNG deficiency on SHM (6). Hence, lack of C-terminus of AID prevents CSR and alters the AID induced mutation pattern in Sμ region.

Also it is important to note that the NES mutant AID (last 17 amino acids deleted) appears to be more catalytically active than full length AID *in vitro* (87), and ΔAID ($AID^{\Delta 189\text{-}198}$) has been shown to be at least as catalytically active as full length AID (34). If this is true *in vivo* also, it is clear that the deamination activity of AID is not adequate to complete CSR, which also requires additional AID-C-terminus-mediated events.

Hypothesis

The lack of the AID C-terminus compromises the efficiency of UNG activity during CSR. Interestingly, RPA32, which interacts with the phosphorylated form of AID, has been shown to interact with UNG by the yeast two-hybrid assay (88). It has been found that *msh6*$^{-/-}$ mice have altered targeting of AID, as the AID induced-mutations in V regions are more focused at WRC hotspots in these mice than in WT mice (28, 30). These facts support the idea that AID might exist in a complex with UNG and MMR proteins. Such a complex might cause AID-induced mutations to more efficiently lead to CSR.

Our results support my hypothesis that AID interacts with UNG and MMR in an AID-C-terminus dependent manner when associated with the Ig Sμ region. This interaction increases the binding of AID, UNG, and Msh2-Msh6 to Sμ, and this interaction is functionally important for CSR.

Our fundamental experimental approach involved isolation of mouse splenic B cells by T cell depletion using anti-T cell antibody and complement. We cultures the B cells for two days in complete medium. We perform retroviral transduction on day two to express AID-ER, ΔAID-ER, or ER. We performed experiments on these cells 24 hours after retroviral transduction (fig 5).

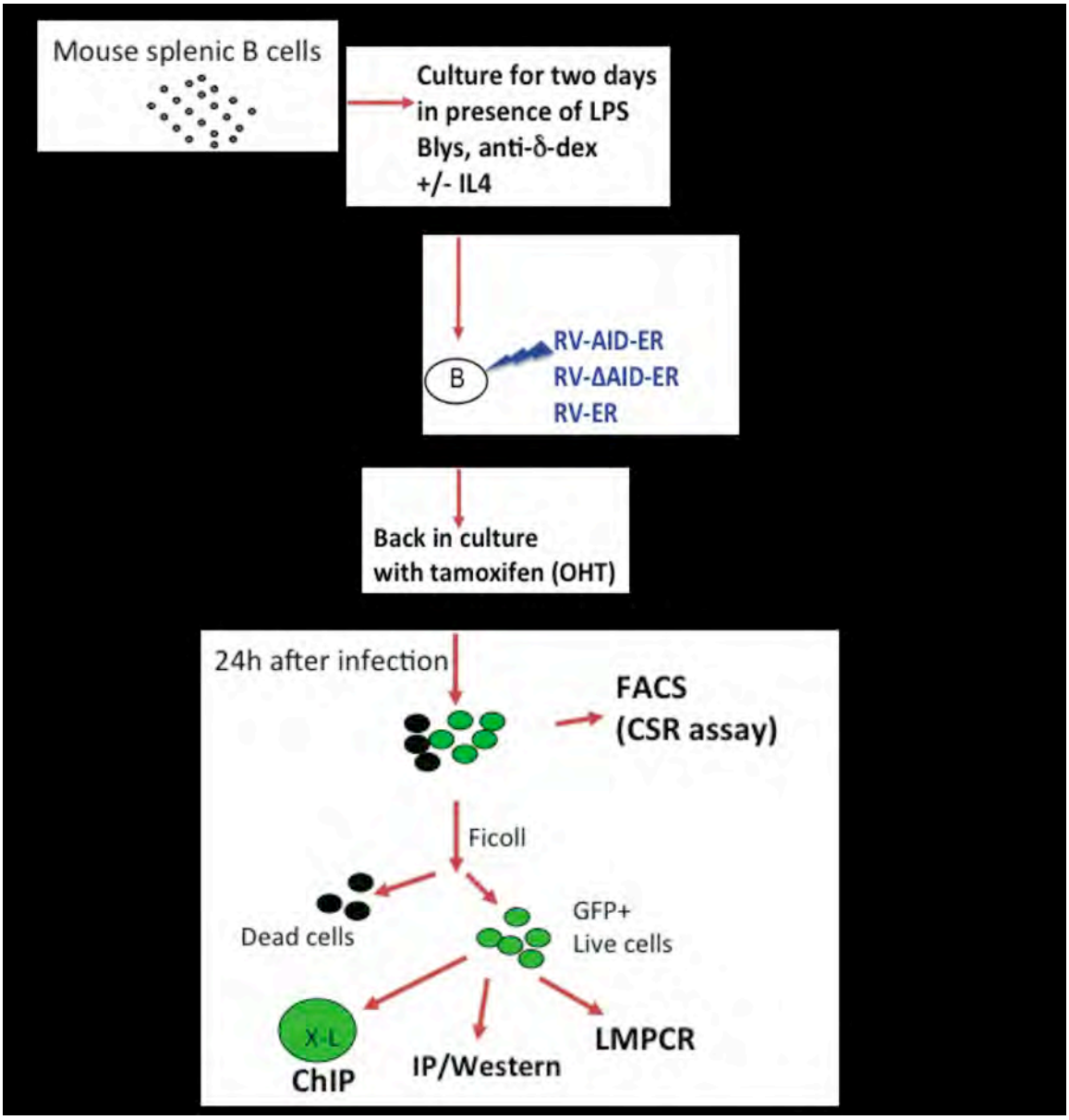
Mouse splenic B cells
Culture for two days
in presence of LPS
Blys, anti-δ-dex
+/- IL4
B
RV-AID-ER
RV-ΔAID-ER
RV-ER
Back in culture
with tamoxifen (OHT)
24h after infection
FACS
(CSR assay)
Ficoll
Dead cells
GFP+
Live cells
X-L
ChIP
IP/Western
LMPCR

Figure 5

Figure 5 Experimental Approach. mouse splenic B cells are purified by T cell depletion using anti-T cell antibody and complement. The B cells were then cultured for two days in a complete medium. We perform retroviral transduction on day two to express AID-ER, ΔAID-ER, or ER in the cultured cells. The cells were then put back in culture in presence of 4OHT for one more day. One day after the retroviral transduction, cells were prepared for FACS analysis or the live cells were purified using Ficoll gradient. The live cells were crossed-linked (X-L) for ChIP analysis, DNA extraction was done to perform LM-PCR, or cell extracts were made for IP and/or western blot.

Chapter II

MATERIALS AND METHODS

Mice. All mouse strains were extensively (eight generations) backcrossed to C75BL/6. AID-deficient mice were obtained from T. Honjo (Kyoto University, Kyoto, Japan) (89). Msh2-deficient mice were obtained from T. Mak, Univ. of Toronto CA (90). Mlh1-deficient mice were obtained from R.M. Liskay, Oregon Health Sciences University, Portland OR (91). UNG-deficient mice were obtained from D. Barnes and T. Lindahl, London Research Institute, UK (92). Msh6-deficient mice were obtained from W. Edelmann (93). Double knockout mice were bred by mating heterozygous mice. Mice were housed in the Institutional Animal Care and Use Committee-approved specific pathogen-free facility at the University of Massachusetts Medical School; these mice were used according to the guidelines from University of Massachusetts Medical School Animal Care and Use committee. For each experiment, splenic B cells were isolated from littermates.

Antibodies. Primary antibodies were for: ER (Santa Cruz Biotechnology sc-8002, lots G2307 and D0110); UNG (rabbit anti-peptide amino acids 280-295 from mouse UNG, prepared by Anaspec Corp); MSH2 (sc-494), MSH6 (sc-10798), and GAPDH (sc-25778). For immunoblotting, secondary antibodies

used were goat-anti-rabbit (sc-2004) or donkey anti-mouse-horseradish peroxidase (sc-2020)(all from Santa Cruz Biotechnology).

Retroviral constructs. pMX-PIE-AID-FLAG-ER-IRES-GFP-*puro* and pMX-PIE-ΔAID-FLAG-ER-IRES-GFP-*puro* (34) were received from Drs V. Barretto and M. Nussenzweig (The Rockefeller University, NY). To create the control retrovirus: pMX-PIE-ER-IRES-GFP, the ER gene from RV-AID-ER was PCR amplified with a forward primer: 5' GCCGGATCCCGCCATGTCTGCTGGAGACATGAGAGCT 3'; and a reverse primer: 5' GGCCTCGAGGAGCTCAAGCTGTGGCAGGGAAACC 3'. The PCR product was digested with BamHI and XhoI, and inserted into BamHI- and XhoI-digested pMX-PIE vector containing IRES, GFP, and puromycin genes downstream of XhoI site. The construct was sequenced and we confirmed proper expression of the RV-ER construct by western blotting.

B cell purification and cultures. Mouse splenic B cells were isolated as previously described, by T cell depletion with antibody and complement (94). B cells were cultured at 10^5 cells/ml in 5% CO2. Lipopolysaccharide (LPS) (50 µg/ml), anti-IgD dextran (α–δ-dex) (0.3 ng/ml) and IL-4 (20 ng/ml) were used to induce class switch recombination to IgG1; LPS and α-δ-dex were used to induce IgG3 switching. Human BAFF/BLyS (50 ng/ml)(Human Genome Sciences) was included in all cultures.

Cell cycle analysis. To analyze cell cycle progression of cells expressing AID-ER, ΔAID -ER, or ER, *aid-/-* mouse splenic B cells were transduced with retroviruses after culture for 48 hours. Cells were harvested one day later, and flow cytometry was used to determine GFP expression levels. Cells were stained with PI and analyzed by flow cytometry to study the cell cycle progression. For this experiment we used cultures in which more than 95% of cells expressed GFP.

Cell proliferation assay. Proliferation of *aid-/-* B cells expressing AID-ER, ΔAID-ER, or ER was assayed by staining freshly isolated *aid-/-* mouse splenic B cells with PKH26 (Sigma), which stains the cell membrane, as prescribed in the product manual. Cells were then cultured for 48 hours, and infected with retroviruses expressing AID-ER, ΔAID -ER, or ER. PKH26 fluorescence was assayed by live cell FACS two days after infection.

Production of retroviruses in Phoenix-E cells.

We seeded $3x10^6$ Phoenix-E cells per 10 cm plate 24 hrs before transfection with the retroviral constructs in 10 ml of complete RPMI1640 medium (10% heat-inactivated serum, 100 µg/ml penicillin- streptomycin, 2 mM L-glutamine, 5 ml non-essential amino-acids (GIBCO 11140), 1 mM sodium pyruvate (GIBCO), 5 ml of 100XBME (β-mercaptoethanol)). 100XBME was made by adding 35µL of

concentrated BME and 1 ml fetal bovine serum in 99ml of 1XPBS. To transfect the retroviral constructs, Fugene 6 (Roche) was brought to room-temp and mixed well before use. 27 µL of Fugene 6 was added to 600 µL of serum free RPMI1640 at room temperature. Fugene 6 was pipetted directly into the medium without contacting the walls of the plastic tube, mixed, and incubated at room-temperature for 5 min. Six µg of retroviral plasmid construct and 3 µg pCL-Eco plasmid (expressing RV-Gag, Pol, and Env proteins) were added, mixed, and incubated for 45 min at room temperature. The mixture was then added drop-wise to the 10 cm plate containing Phoenix-E cells, and the culture dish was swirled to ensure even distribution of the mixture. The plate was incubated at 37°C for 24 hrs, and moved to a 32°C incubator, which resulted in a higher titer of virus production (the virus is unstable at higher temperatures). Twenty-four hours after the transfection, the transfection efficiency of Phoenix cells was checked by FACS for GFP expression. Cultures that have ~65-90% GFP^+ cells yielded good to excellent virus titers. Virus supernatant was collected 2 days after transfection, and replaced with 10 ml of fresh complete RPMI1640 with heat-inactivated serum. The supernatant was then centrifuged at 1200 rpm for 5 min. Virus was harvested 2 more times every 48 hrs thereafter. The supernatant was used immediately or snap-frozen in liquid nitrogen before storing at -80°C. Since the virus-titer drops by 50% with each freeze-thaw cycle, appropriate aliquots were made before they were frozen.

Retroviral infection of B cells. 4 ml of B cells at 10^5 per ml were cultured for two days in 6 well plates or 2 ml at 10^5 per ml in 24-well plates in IgG1 (LPS+Blys+ anti-δ-dextran+ IL4) or IgG3 (LPS+Blys+ anti-δ-dextran) CSR conditions. Each culture was split into two cultures on day 2. The split culture plates were centrifuged at 1200 rpm for 5 min after the split in order to remove the supernatant without removing the cells. The cells on the plate were re-suspended in the retrovirus cocktail containing: virus supernatant 1 ml, RPMI1640 1ml, heat-inactivated FBS 30 µL, LPS (50 µg/ml), IL4 (20 ng/ml) (not included in IgG3 switching conditions), BLyS (50 ng/ml), anti-δ-dex (0.3 ng/ml), 4OHT (Sigma) (1 µM) (34) and polybrene (10 µg/mL) (34). After incubating the cocktail for 45 min at room temp, the cells were resuspended in the cocktail and the culture plates were centrifuged at 2000 rpm (Beckman-Coulter, Allegra-12R centrifuge) for 45 minutes at 20°C. The plates were returned to 37°C, 5% CO_2 for 24 hrs, and flow-cytometric analysis was done to check GFP expression level and CSR assay.

CSR Assay and Flow-cytometric analysis. To study CSR in B cells expressing AID-ER, ΔAID-ER, or ER, we cultured 2 ml of B cells (10^5 per ml) for 2 days in IgG1 or IgG3 switching conditions described above, and infected these cells with retrovirus to express AID-ER, ΔAID-ER, or ER. The cells were treated with 1 µM 4OHT at the time of infection. Cells were harvested 24 hr after infection and subjected to (Fluorescence Activated Cell Sorting) FACS staining

and analysis was partly done as described previously (94), except that cells were also treated with 7-AAD to eliminate dead and dying cells more efficiently. Twenty-four hours after the retroviral transduction, cells were washed three times with 2 ml of FACS buffer (1XPBS with 2% FBS), each time cells were pelleted by centrifuging the tubes at 335Xg at 4°C. Cells were kept in ice at all times. After the last wash, cells were resuspended in 50 µL of the FACS buffer. 0.0025 µg/ml PE-IgG1 (Southern Biotech) or 0.001 µg/ml PE-IgG3 (Southern Biotech) were added to the cell suspension and the tubes were incubated in ice for 30 minutes. The stained cells were washed again 3 times with the FACS buffer, and were suspended in 400 µL of the FACS buffer after the final wash. 0.8 µg/µl 7-AAD (Imgenex) was added to the final cell suspension.

Flow-cytometric analysis was performed using Flowjo software, first 7-AAD negative cells were selected to eliminate the dead or dying cells. GFP-positive cells were gated within the 7-AAD negative cells. Then PE-IgG1 or PE-IgG3 cells were gated within the 7-AAD-negative and GFP-positive cells.

Chromatin immunoprecipitation (ChIP). ChIP assays were performed as described previously (95). B cells were cultured for 2 days in IgG3 conditions, retroviral transduction was done on day two. The cells were treated with 4OHT at the time of retroviral transduction. Live cells were isolated by flotation on Lympholyte M (Cedar Lane, Ontario Canada) 24 hours after retroviral

transduction. 1 ml of Lympholyte M at room temperature was placed in 15 ml Falcon tubes (BD), and overlayed with 8 ml of the cell culture. The tubes were then centrifuged at 1200Xg for 20 minutes at room temperature. The live cells, which float on the top layer of the Lympholyte M were gently removed by a pasture-pipette. The live cells were then washed twice in serum free 1XPBS. After the 2^{nd} wash, $2X10^7$ cells were resuspended in 1XPBS and were cross-linked with formaldehyde at a final concentration of 1% for 2 min at 37°C. Then cross-linking was stopped by adding 125 mM glycine (final concentration) and incubating at room temperature for 5 minutes.

The cross-linked cells were washed twice with 1XPBS then lysed in ice for 10 minutes in lysis buffer (1%SDS, 10 mM EDTA, 50 mM Tri-Cl pH 8.1, 1X EDTA–free protease inhibitor cocktail (Roche), 10 mM NaF (sodium fluoride), 1 mM Na3VO4 (sodiumorthovanadate), 10 µg/ml PMSF (phenylmethanesulfonyl-fluoride) in distilled water, then sonicated on ice for 3 minutes. Sonication was performed in 10 sec bursts, the samples were incubated for 15 seconds in ice between each bursts. The cell lysates were precleared with 75 µL 50% pre coated protein A/G-agarose beads, rotated end-to-end at 4°C for 2 hours. 10^6 cell equivalents of precleared cell lysate were used per IP, and 10^5 cells were used for the input samples. To correct for this difference, ChIP results were divided by 10. Significance was calculated by a two-tailed T test. ChIP results were assayed by real time PCR using Sybr Green (Invitrogen). Primers for Sµ were: (forward)

DK99 5'-AAC TAG GCT GGC TTA ACC GAG ATG-3', (reverse) DK100 5'-GTC CAG TGT AGG CAG TAG AGT TTA-3'; Primers for Cμ were: (forward) DK46 5'-GTC AGT CCT TCC CAA ATG TCT TCC-3', (reverse) DK47 5'-CTG GAA TGG GCA CAT GCA GAT CTT T-3'.

For immunoprecipitation, 25 µl of pre-cleared cell lysate containing 10^6 cell equivalents were added to 1.4 ml dilution buffer. The dilution buffer contained distilled water with 0.01% SDS, 1.1% Triton-X 100, 1.2 mM EDTA, 16.7 mM Tris-Cl pH 8.1, 167 mM NaCl, 1X EDTA–free protease inhibitor cocktail, 10 mM NaF, 1 mM Na3VO4 , 10 µg/ml PMSF. Immunoprecipitation was done with 2-8 µg of antibody, which was added to the diluted cell lysate and incubated (end-to-end rotation) at 4°C over night. On the following day, 50 µL 50% coated protein A-G beads were added to each tube and incubated at 4°C for 2 hours with end-to-end rotation. Then the beads were pelleted by centrifuging at 500Xg. The beads were then washed 5 minutes each (end-to-end rotation at 4°C) with low-salt wash buffer (0.1% SDS, 1% Triton-X 100, 2 mM EDTA, 20 mM Tris-Cl pH 8.1, 150 mM NaCl, 1X EDTA–free protease inhibitor cocktail, 10 mM NaF, 1 mM Na3VO4 , 10 µg/ml PMSF in distilled water), high-salt wash buffer (0.1% SDS, 1% Triton-X 100, 2 mM EDTA, 20 mM Tris-Cl pH 8.1, 500 mM NaCl, 1X EDTA–free protease inhibitor cocktail, 10 mM NaF, 1 mM Na3VO4 , 10 µg/ml PMSF in distilled water), Lithium-salt wash buffer (250 mM LiCl, 1% Deoxycholate, 1% Nonidet P-40, 1 mM EDTA, 10 mM Tris-Cl pH 8.1 in distilled water), and TE (10 mM Tris-Cl pH

8.1, 1 mM EDTA, 1X EDTA–free protease inhibitor cocktail, 10 mM NaF, 1 mM Na3VO4 , 10 µg/ml PMSF). After the final wash, the excess TE was removed as much as possible. The beads were then resuspended in 400 µl of elution buffer (10 mM Tris-Cl pH 8.1, 5 mM EDTA, 300 mM NaCl, 0.5% SDS in distilled water), and incubated at 65°C over-night. Then the temperature was reduced to 55°C, 2 µL 10 mg/ml proteinase K was added and incubated again over-night. The DNA was purified by 2 phenol-chloroform (1:1) extraction followed by ethanol precipitation (1 µL of 20 mg/ml glycogen in 1 ml 95% ethanol) over-night at -20°C. The precipitated DNA was washed twice with 75% ethanol, and the DNA pellet was dried in air. The dry pellet was then dissolved in 25 µL of the TE buffer without protease inhibitors. The DNA was then analyzed by a real-time PCR. The input DNA by prepared by adding 10 µl of the cross-linked cell lysates to 400 µl TE without protease inhibitors. 30 µl 5 M NaCl was added and uncross-linked over-night at 65°C. Then 1 µl of 10 mg/ml RNase and 2 µl 10 mg/ml Proteinase K were added and incubated at 55°C over-night. Then ethanol purification of the DNA was done as described above. Dry input DNA pellet was resuspended in 75 µl of the TE buffer without protease inhibitors.

In vitro transcription translation. The T7 promoter was fused to the AID-ER, ΔAID-ER, ER, GST and GST-UNG genes. The constructs were then inserted into pCSR®4-TOPO (Invitrogen) plasmid. In vitro transcription and translation were performed using the PROTEINscript® II T7 Kit, as prescribed in the

manufacturer's manual (Ambion). *In vitro* transcription and translation was optimized for comparable levels of expression of AID-ER, ΔAID-ER, and ER also, GST and GST-UNG.

GST Pull-down. Equal volume of GST or GST-UNG was added to each tube containing AID-ER, ΔAID -ER and ER. Dilution buffer containing: 1XPBS with Complete protease inhibitor cocktail (Roche; 25X=1 tab/2 ml 1XPBS), 200 µM phenylmethanesulfonylfluoride (PMSF), 200 µM sodium orthovanadate, 50 µM NaF, 2 µM β-glycerophosphate, 50 µM $ZnCl_2$, 200 µg/ml ethidium bromide was added up to the final volume of 500 µL. The mixtures were incubated at 4°C for 12 hrs and then incubated with glutathione beads for another 2 hrs. The beads were spun down and then washed 3X with wash buffer. The wash buffer was made by adding 200 mM NaCl and 0.05% NP40 to the dilution buffer without ethidium bromide. The proteins were then eluted with elution buffer after the third wash. The eluates were then subjected to western blot.

Genomic DNA preparation and linker ligation-mediated PCR (LM-PCR). After culture for two days and 24 hours after retroviral transduction, viable cells were isolated by flotation on Lympholyte M (Cedar Lane, Ontario, Canada), cells were imbedded in low melt agarose plugs, DNA isolated and LM-PCR performed as described (42), with slight modifications. For linker ligation, 50 µL 1X ligase buffer was added to the plugs, which were then heated to 62°C to melt the

agarose. 20 µL DNA (about 200,000 cell equivalents) was added to 2 µL T4 DNA Ligase (2 Weiss units, MBI Fermentas, Hanover, MD), 10 µL ds annealed linker in 1X ligase buffer, 3 µL 10x Ligase buffer and 30 µL dH_20 and incubated overnight at 18°C. Linker was prepared by annealing 5 nmoles each of LMPCR.1 (5'-GCGGTGACCCGGGAGATCTGAATTC-3') and LMPCR.2 (5'-GAATTCAGATC-3') in 300 µL 1X ligase buffer, which results in a ds oligo with a 14 nt ss overhang that can only ligate unidirectionally. Ligated DNA samples were heated at 70°C for 10 min, diluted 5X in dH_20 and then assayed for *gapdh* DNA by PCR to adjust DNA input prior to LM-PCR. The primers 5'Sµ (5'-GCAGAAAATTTAGATAAAATGGATACCTCAGTGG-3'), Sg3-AP: 5′-AACATTTCCAGGGACCCCGGAGGAG-3′; and CmuL2: 5′-CTGCGAGAGCCCCCTGTCTGATAAG-3′ were used in conjunction with linker primer (LMPCR.1) to amplify DNA breaks in Sµ, Sγ3, or Cµ respectively. Three-fold dilutions of input DNA (0.5, 1.5 and 4.5 µL for Sµ/Cµ LM-PCR, 1.5, 4.5 and 13.5 µL for Sγ3 LM-PCR) were amplified by HotStar Taq (Qiagen) using a touchdown PCR program (28 cycles after touchdown for Sµ LM-PCR and 35 cycles after touchdown for Sγ3 LM-PCR). PCR products were electrophoresed on 1.25% agarose gels and blotted onto nylon membranes (GeneScreen Plus, Perkin Elmer, Waltham, MA). Blots were hybridized with an Sµ-specific oligonucleotide probe (mu probe5': 5'-AGGGACCCAGGCTAAGAAGGCAAT-3') for 5' Sµ LM-PCR, or Sγ3-LP: 5′-GGACCCCGGAGGAGTTTCCATGATCCTGGG-3′ or Cµ: 5′-

TGGCCATGGGCTGCCTAGCCCGGGACTTCCTG-3′ that had been end-labeled with [γ^{32}P]-ATP. Semi-quantitative assessment of DSB frequency was achieved by densitometry scanning of autoradiographs, combining all 3 dose titration lanes for each set (from one mouse). For determination of the sites of DSBs, PCR products were cloned into pCR4-TOPO (Invitrogen, Carlsbad, CA).

Immunoprecipitation and western blot. Retrovirus transduced live B cells were purified by Lympholyte M floatation. The live cells were then subjected to whole cell extraction. To make the whole cell extract, 10^7 cells were incubated in 100 µl RIPA buffer (25 mM Tris•HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS, 10 mM Tris-Cl pH 8.1, 1 mM EDTA, 1X EDTA–free protease inhibitor cocktail, 10 mM NaF, 1 mM Na3VO4, 10 µg/ml PMSF) in ice for 5 minutes. The cell lysate was passed through a pipette 15-20 times to make a homogeneous mixture. The mixture was again incubated in ice for 5 minutes, then centrifuged at 14,000xg for 15 minutes at 4°C to separate the cell debris from the protein extracts. The protein extracts were then quantified by Bradford assay.

For immunoprecipitation, the cell lysates were precleared with 75µl of 50% coated protein A/G slurry. The lysates were diluted to 1.4 ml using dilution buffer (0.01% SDS, 1.1% Triton-X 100, 1.2 mM EDTA, 16.7 mM Tris-Cl pH 8.1, 167 mM NaCl, 1X EDTA–free protease inhibitor cocktail, 10 mM NaF, 1 mM Na3VO4,

10 µg/ml PMSF in distilled water). 2-4 µg of antibodies were added and incubated overnight at 4°C. Non-specific IgG was used as negative control. The following day, the beads were pelleted by centrifuging at 500Xg. The supernatant was carefully removed and discarded without disturbing the bead pellet. The beads were washed once with 1XPBS containing 1X EDTA–free protease inhibitor cocktail, 10 mM NaF, 1 mM Na3VO4, 10 µg/ml PMSF, twice with 1XPBS containing 150 mM NaCL, 0.1% NP-40, 1X EDTA–free protease inhibitor cocktail, 10 mM NaF, 1 mM Na3VO4, 10 µg/ml PMSF. After the third wash, the beads were resuspended and boiled for 5 minutes in 50 µL of loading buffer containing 300 mM of NaCl. Then the solution with beads was allowed to cool to room temperature, centrifuged at max speed for 2 minutes and the supernatant was loaded onto 5-20% gradient protein gel. Electrophoresis was performed in 1XHepes buffer (0.1 M Tris, 0.1 M HEPES, 0.1% SDS, pH 8).

After electrophoresis, the protein samples in the gel were transferred onto Immobilon-P transfer membrane (size 0.45 µm). Blots were blocked with 5% non-fat milk for two hours, then incubated with 10 ml of the milk containing 2-4 µg of antibodies overnight. The blot is then washed 3 times with 1X wash buffer (1 L 1XPBS with 250 µL NP-40). The washed blot is incubated in 10 ml 5% milk containing secondary antibodies for 2 hours at room-temperature. The blot was washed 3 times with the 1X wash buffer, and then exposed to X-ray film.

Chapter III

RESULTS

Expression of ΔAID-ER or AID-ER does not alter cell cycle progression or cell proliferation.

To test the role of the AID C terminus in CSR, we expressed full length AID or truncated AID lacking the C terminal 10 amino acids by retroviral transduction in activated *aid-/-* mouse splenic B cells. The retroviral constructs encode full-length AID or ΔAID (Δ189-198) fused to a FLAG tag and the estrogen receptor (ER) (AID-ER and ΔAID-ER, respectively) (34) (Fig 6a). The fusion protein is followed by an internal ribosome entry site (IRES) and a gene encoding enhanced green fluorescent protein (EGFP). This entire construct is flanked by long terminal repeats (LTR). We also prepared a control vector that expressed the ER protein but no AID.

The three retroviral constructs (RV-AID-ER, RV-ΔAID-ER or RV-ER) were transduced into *aid-/-* splenic B cells two days after activation with LPS, BlyS, and anti-IgD dextran. The cultures were treated with tamoxifen (4OHT) at the time of infection, which induces nuclear localization of ER-tagged proteins, in order to allow AID to reach its target. One day later, transduction efficiency was assayed by GFP expression (Fig 6b). GFP expression in the experiments presented in this study varied from ~65-95%. In the studies reported here, we

only include experiments in which the three different retroviruses yielded similar GFP expression. Consistent with previous reports, full length AID-ER induces ~10-fold more IgG1 and IgG3 CSR than ΔAID-ER in *aid-/-* splenic B cells, as shown in the flow cytometry analyses in Fig 6f. In preliminary experiments, we found that nearly all the AID-ER protein is in the nucleus and that the expression levels of AID-ER and ΔAID-ER proteins in cells with similar GFP expression are quite similar (Fig 6c). Previously, it was reported that ΔAID-ER is not expressed as well as AID-ER in splenic B cells (34). Perhaps this is because in those experiments, cells were cultured for 3 days after viral transduction, whereas we assay our experiments one day after transduction. We chose this time because there is less cell death on day 1 than on subsequent days.

Since CSR is cell division linked (96), it is important to examine whether the cell cycle profiles and cell proliferation rates in cells expressing ΔAID-ER are similar to that of cells expressing AID-ER. We stained cells expressing AID-ER, ΔAID-ER, or ER with propidium iodide (PI) on day 1 after transduction and examined the cell cycle profile of each using flow cytometry. AID-ER or ΔAID-ER expression has no effect on cell cycle progression of *aid-/-* B cells (Fig 6d). To examine cell proliferation, B cells were stained with PKH26, which binds to the cell membrane, prior to activation and transduction with RV-AID-ER, RV-ΔAID-ER, or RV-ER. Flow cytometry indicated that cells transduced with the three different RV constructs proliferate similarly (Fig 6e). We also performed CSR

assay to make sure that our retroviral constructs behaved similarly to previously published results, which showed that CSR was reduced by 80-90% in ΔAID-ER expressing cells (Fig 6f). Western blots of nuclear extracts of B cells expressing AID-ER, ΔAID-ER, and ER showed similar nuclear localization of the proteins in cells treated with 4OHT at the time of retroviral transduction (Fig 6g).

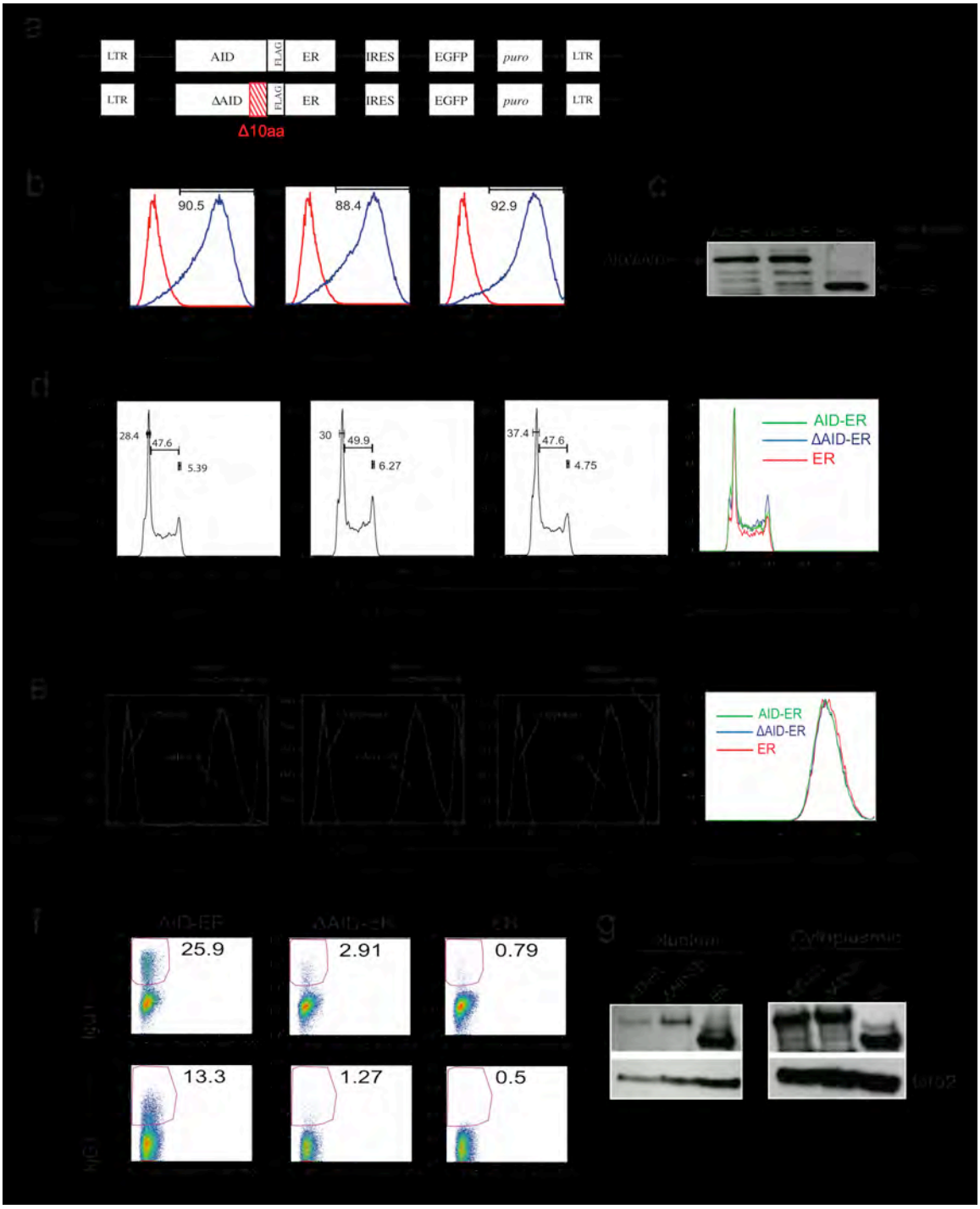
LTR
AID
FLAG
ER
IRES
EGFP
puro
LTR
LTR
ΔAID
FLAG
ER
IRES
EGFP
puro
LTR
Δ10aa
90.5
88.4
92.9
28.4
47.6
5.39
30
49.9
6.27
37.4
47.6
4.75
AID-ER
ΔAID-ER
ER
AID-ER
ΔAID-ER
ER
25.9
2.91
0.79
13.3
1.27
0.5

Figure 6

Figure 6 Transduction of RV-AID-ER, RV-ΔAID-ER, and RV-ER in aid-/- mouse splenic B cells does not alter cell cycle progression or cell proliferation. (a) RV-AID and RV-ΔAID constructs each contain an AID-FLAG-ER fusion protein gene, an internal ribosome entry site (IRES), a gene encoding enhanced green fluorescent protein (EGFP), and a puromycin selection marker. These genes are flanked by RV Long Terminal Repeat (LTR) sequences. RV-ΔAID has amino acids 189-198 deleted. (b) Examples of FACS profiles to show EGFP (GFP) expression in aid-/- mouse splenic B cells one day after transduction with RV-AID-ER, RV-ΔAID-ER, or RV-ER. (c) Western blot of whole cell extracts (WCE) of aid-/- B cells show equivalent expression levels of RV-AID-ER, RV-ΔAID-ER, and RV-ER in extracts from equal numbers of transduced cells. On the gels, the apparent molecular weight of the AID-ER protein is ~90 kD. (d) Cell cycle profiles of aid-/- B cells one day after transduction with RV-AID-ER, RV-ΔAID-ER, or RV-ER; overlap of the profiles (right most panel) demonstrates similar cell cycle profiles. (e) Cell proliferation profile of aid-/- B cells which were stained with PKH26 on day 0, activated for 2 days, then transduced with RV-AID-ER, RV-ΔAID-ER, or RV-ER, and harvested 2 days after transduction. Overlapped profiles (right most panel) demonstrate identical proliferation. (f) FACS profiles of CSR assay showing switching to IgG1/IgG3 in aid-/- B cells expressing AID-ER, ΔAID-ER, or ER. Each panel shows viable and GFP^+ cells, each gate represents PE-IgG1/PE-IgG3 positive cells, as indicated within the GFP^+ population. The profiles show ~90% reduction in CSR in ΔAID-ER compared to that of the AID-ER expressing cells. **(g)** western blot of Nuclear and cytoplasmic fraction of *aid-/-* B cells expressing AID-ER, ΔAID-ER, or ER show similar levels of AID-ER, ΔAID-ER and ER proteins in the nucleus and cytoplasm of cells treated with 4OHT at the time of retroviral transduction.

AID and UNG bind co-dependently to Sμ DNA dependent upon the AID C-terminus.

To begin to explore the function of the C terminus of AID, we asked if ΔAID-ER associates with the Sμ region in vivo. ΔAID-ER has been shown to cause as many mutations as AID-ER in the 5' region of Sμ in cells induced to undergo CSR (34). Also, ΔAID has been found to have equivalent or higher catalytic activity than full-length AID in cell-free experiments (87, 97). We expressed AID-ER, ΔAID-ER, or ER alone in *aid-/-* splenic B cells, and performed chromatin immunoprecipitation (ChIP) using antibody for ER. Surprisingly, the ChIP data show decreased association of ΔAID-ER with Sμ DNA compared to AID-ER (Fig 7a). As anti-ER antibody can precipitate AID-ER, ΔAID-ER, and ER proteins equally well (Fig 7b), we conclude that binding of AID-ER to the Sμ region that is stable enough to be detected by ChIP depends upon its C terminus. Since in the absence of the C terminus, AID can still induce mutations at Sμ, it is clear that this stable association with Sμ is not important for its ability to deaminate Sμ, although it might be important for CSR.

Cells expressing ΔAID-ER show an increased proportion of C>T and G>A transition mutations relative to C>A and C>G or G>C and G>T transversion mutations (34), similar to the phenotype of *ung-/-* B cells. This observation suggests that AID might associate with UNG via the AID C terminus, and help recruit UNG to switch regions. We asked, therefore, if association of

endogenous UNG with Sμ is compromised in cells expressing the C terminal deleted AID. Indeed, ChIP with anti-UNG antibody demonstrated that whereas UNG associates with Sμ in *aid-/-* B cells expressing AID-ER, it is found only at background levels in cells expressing ΔAID-ER (Fig 7d). Neither AID-ER, ΔAID-ER nor UNG are found at the Cμ region in these cells, which indicates that the association is specific for Sμ (Fig 7a,d).

Our finding that binding of UNG to Sμ in vivo depends on the AID C terminus, prompted us to ask if the preferential binding of full-length AID to Sμ likewise depends upon UNG. We expressed AID-ER, ΔAID-ER, and ER in *aid-/- ung-/-* splenic B cells and performed ChIP with anti-ER antibody. Our results clearly show that in cells lacking UNG, full length AID-ER does not associate better than ΔAID-ER with Sμ (Fig 7a). Therefore, the ability of full-length AID but not C terminal deleted AID to associate with Sμ depends upon the presence of UNG. These data suggest that AID and UNG bind co-dependently to the Sμ region in B cells undergoing CSR, and that this co-dependently binding depends on the AID C-terminus.

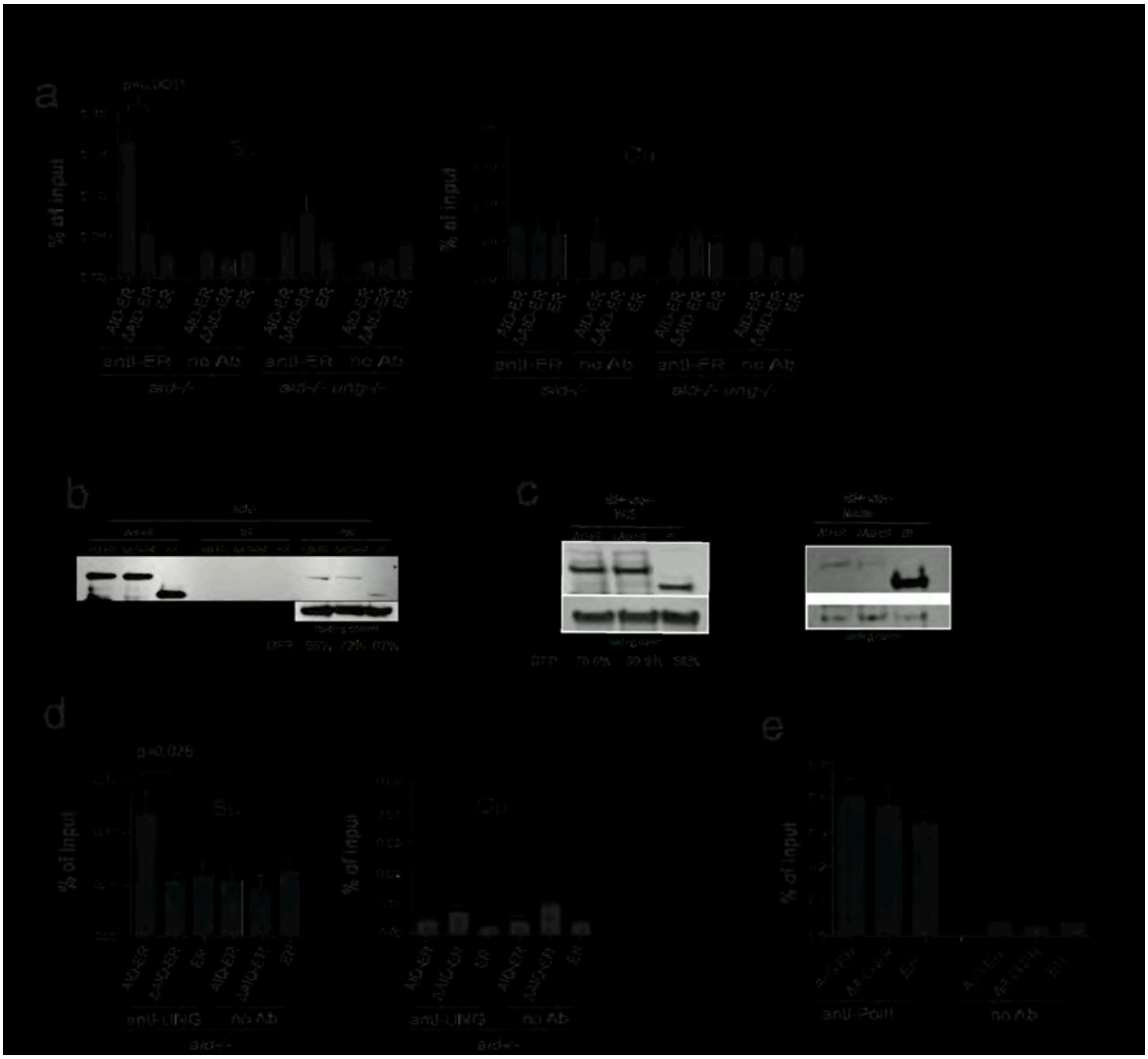
a
% of input
anti-ER
no Ab
b
c
d
% of input
anti-UNG
no Ab
e
% of input
anti-Pol
no Ab

Figure 7

Figure 7 Chromatin immunoprecipitation (ChIP) demonstrates that the C-terminus of AID is required for recruitment of AID-ER and UNG to Sμ DNA, and that UNG is important for recruitment of AID-ER to Sμ. (a) Anti-ER ChIP of extracts from *aid-/-* splenic B cells expressing AID-ER, ΔAID-ER, or ER shows association of AID-ER with Sμ in *aid-/-* B cells but not in *aid-/- ung-/-* cells, and not at Cμ in either genotype. ChIP was analyzed by real-time PCR. Means +SEM of 3 or more ChIPs for Sμ and Cμ are shown. (**b)** Western blot of IP performed using anti-ER antibody and normal IgG (negative control) on WCE of *aid-/-* splenic B cells expressing AID-ER, RV-ΔAID-ER, or RV-ER shows equivalent expression levels of RV-AID-ER, RV-ΔAID-ER and RV-ER. Msh2 is the loading control in b. Percent of cells, which express GFP in each culture, is shown at the bottom of the corresponding lane. **(c)** Western-blot of WCE of *aid-/- ung-/-* splenic B cells expressing RV-AID-ER, RV-ΔAID-ER, or RV-ER. GAPDH is the loading control (left panel). Western-blot of nuclear extract of *aid-/- ung-/-* splenic B cells expressing RV-AID-ER, RV-ΔAID-ER, or RV-ER. Non-specific bands are the loading control (right panel) **(d)** Anti-UNG ChIP of extracts from *aid-/-* splenic B cells expressing RV-AID-ER, RV-ΔAID-ER, or RV-ER, showing specific association of UNG with Sμ in RV-AID-ER but not RV-ΔAID-ER expressing *aid-/-* splenic B cells. Means+SEM of 3 or more ChIPs for Sμ and Cμ are shown. **(e)** Anti-PolII ChIP shows specific association of PolII with Sμ in RV-AID-ER, RV-ΔAID-ER, and ER expressing *aid-/-* splenic B cells. The differences among the three are not significant. Means+SEM of 3 ChIPs for Sμ and Cμ are shown.

AID and Msh2-Msh6 bind co-dependently with Sμ DNA dependent upon the AID C-terminus.

The mismatch repair (MMR) heterodimer Msh2-Msh6 binds to U:G mismatches that result from AID activity (98), and the specificity of AID-induced mutations at G:C bp is altered in *msh6-/-* cells (28, 30). We asked if association of endogenous Msh2-Msh6 with Sμ in vivo depends upon the C terminus of AID. ChIP for Msh2 and for Msh6 in *aid-/-* B cells expressing AID-ER, ΔAID-ER, or ER demonstrated that both Msh2 and Msh6 associate with Sμ in cells expressing full-length AID-ER, but not significantly above background when expressing ΔAID-ER or ER (Fig 8a). Binding of Msh2 and Msh6 is not detected at the Cμ region (Fig 8b), in agreement with our finding that AID-ER binds with Sμ and not Cμ.

We next asked if binding of full-length AID to Sμ in vivo might depend upon these MMR proteins by expressing the RV constructs in *aid*$^{-/-}$*msh2*$^{-/-}$ and in *aid*$^{-/-}$*msh6*$^{-/-}$ cells. Indeed, we found that in the absence of Msh2 or Msh6, AID-ER does not bind better than ΔAID-ER to the Sμ region, and the binding was at background levels (Fig 8c). In fact, we detected no binding above background of AID-ER in *aid-/-msh2-/-* cells. This is not due to differential expression of the AID proteins, as the levels of AID-ER and ΔAID-ER are similar in B cells from both *aid-/- msh2-/-* and *aid-/- msh6-/-* mice (Fig 8d). Together with the results showing reduced

binding of Msh2-Msh6 to Sμ in cells expressing ΔAID-ER, these data suggest that AID binds co-dependently with Msh2-Msh6 to the Sμ region and that this co-dependent binding depends on the C terminus of AID. To reconcile these results with published results indicating that AID and ΔAID have similar abilities to generate mutations in the Sμ region (34), we hypothesize that co-dependent binding of AID with UNG and Msh2-Msh6 stabilizes the binding of AID to Sμ, thus allowing the binding to be detected by ChIP. This increased stability does not appear necessary, however, for the deamination activity of AID.

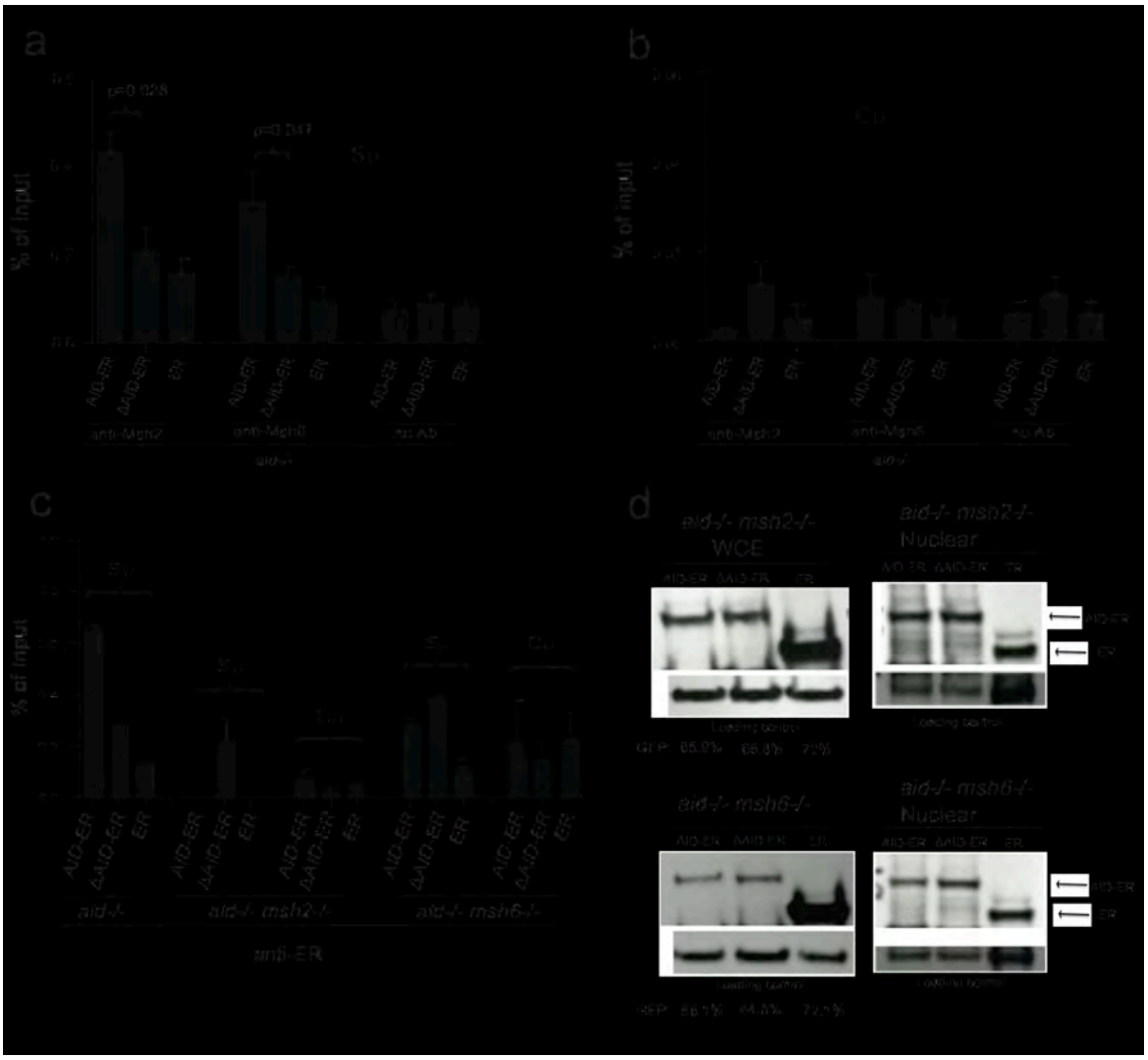
a
b
c
d
% of input
AID-ER
ΔAID-ER
ER
anti-Msh2
anti-Msh6
aid-/-
aid-/- msh2-/-
aid-/- msh6-/-
anti-ER
aid-/- msh2-/-
WCE
aid-/- msh2-/-
Nuclear
aid-/- msh6-/-
WCE
aid-/- msh6-/-
Nuclear

Figure 8

Figure 8 ChIP demonstrates that the C-terminus of AID is important for recruitment of Msh2-Msh6 to Sμ DNA, and that Msh2-Msh6 is important for binding of AID-ER to Sμ. (a) Anti-Msh2 and anti-Msh6 ChIP of extracts from *aid-/-* splenic B cells expressing RV-AID-ER, RV-ΔAID-ER, or RV-ER shows association of Msh2 and Msh6 with Sμ, but not Cμ (b), in *aid-/-* B cells expressing RV-AID-ER but not in cells expressing RV-ΔAID-ER. Means+SEM of 3 or more ChIPs for Sμ are shown. **(c)** Anti-ER ChIP of extracts from *aid-/-*, *aid-/-msh2-/-* and *aid-/-msh6-/-* splenic B cells expressing RV-AID-ER, RV-ΔAID-ER, or RV-ER. One ChIP was done as a positive control in *aid-/-* B cells. Means of two or more ChIPs for Sμ and Cμ are shown for *aid-/-msh2-/-* and *aid-/-msh6-/-* B cells. Error bars indicate either the range (when 2 ChIPs performed). **(d)** Western blots of WCE (top left panel) and nuclear extract (top right panel) from *aid-/-msh2-/-* and *aid-/-msh6-/-* (bottom left and bottom right panel respectively) mouse splenic B cells expressing RV-AID-ER, RV-ΔAID-ER, or RV-ER, show equivalent levels of expression of RV-AID-ER and RV-ΔAID-ER as well as equivalent levels of nuclear localization in the cells treated with 4OHT at the time of retroviral transduction. GAPDH is the loading control in WCE and non-specific band in nuclear extract.

The C terminus of AID is required for the function of MMR proteins during CSR.

CSR is reduced by ~50-70% in MMR-deficient B cells (1, 27, 28, 94, 99). As we found that the co-dependent binding between AID and the Msh2-Msh6 heterodimer at Sμ depends upon the AID C terminus, we next asked if the C terminus of AID is important for the function of MMR during CSR. Although ΔAID deaminates dC's in Sμ, thereby producing the U:G substrate for MMR, we hypothesized that recruitment of Msh2-Msh6 by AID is essential for CSR, and that in cells expressing ΔAID-ER, deficiencies in Msh2 or Msh6 proteins will not reduce CSR efficiency relative to MMR-sufficient cells. To test this hypothesis, we examined CSR to IgG1 and IgG3 in *aid-/-msh2-/-* and in *aid-/-msh6-/-* B cells transduced with the three RV constructs (Fig 9). In cells expressing full-length AID, CSR was reduced by 45-70% by Msh2 or Msh6-deficiency, consistent with results in MMR-deficient cells expressing endogenous AID. Although switching to IgG1 in *aid-/-* B cells expressing ΔAID-ER is reduced by ~90% compared to that of cells expressing AID-ER, it is not further reduced by deficiencies in Msh2 (Fig 9a,b, upper panels) or Msh6 (Fig 9c,d, upper panels). Similarly, switching to IgG3 in ΔAID-ER-expressing *aid-/-* B cells is reduced by ~90% relative to AID-ER-expressing *aid-/-* B cells, but not further reduced in *aid-/- msh2-/-* B cells (Fig 9a,b, lower panels) or in *aid-/- msh6-/-* B cells (Fig 9c,d, lower panels).

We conclude from these data that the contributions of Msh2 and Msh6 to CSR are abolished in B cells expressing ΔAID-ER. As C terminal deleted AID is an active deaminase and induces Sμ mutations, this is not due to the lack of substrates for Msh2-Msh6 in Sμ in cells expressing ΔAID-ER. These results also indicate that despite being over-expressed, AID-ER behaves similarly to endogenous AID as optimal CSR depends upon MMR. Most interestingly, the data suggest that the ability of the AID C terminus to directly or indirectly recruit MMR proteins to the Sμ region is important for the function of MMR during CSR.

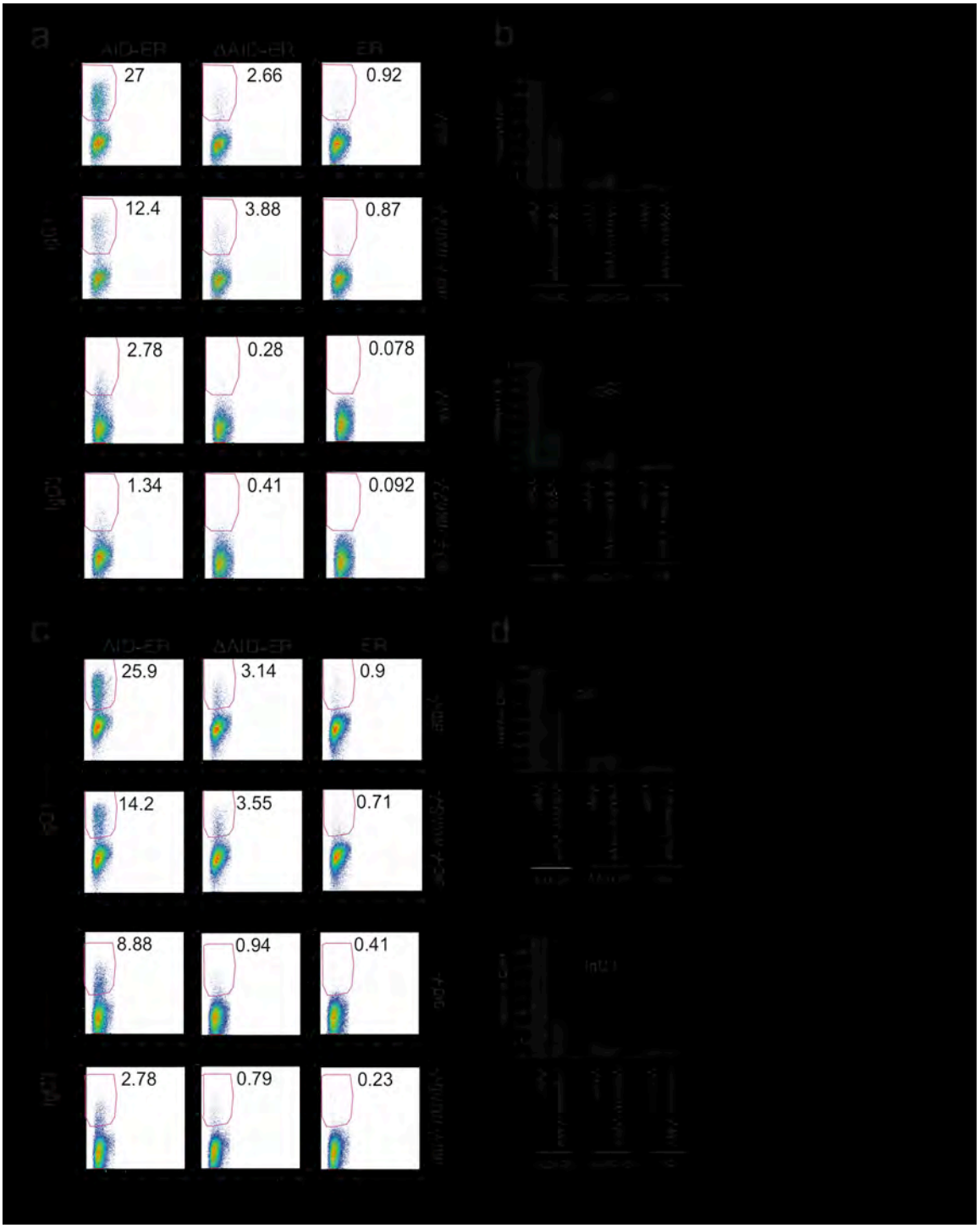
a
AID-ER
ΔAID-ER
ER
27
2.66
0.92
12.4
3.88
0.87
2.78
0.28
0.078
1.34
0.41
0.092
b
c
AID-ER
ΔAID-ER
ER
25.9
3.14
0.9
14.2
3.55
0.71
8.88
0.94
0.41
2.78
0.79
0.23
d

Figure 9

Figure 9 Msh2-Msh6 proteins increase CSR in cells expressing AID-ER but not in cells expressing ΔAID-ER. **(a)** FACS results for one experiment showing CSR to IgG1 and IgG3 (indicated on left) in *aid-/-* and *aid-/-msh2-/-* splenic B cells expressing RV-AID-ER, RV-ΔAID-ER, or RV-ER. Each panel shows viable cells that are GFP positive, and each gate shows PE-IgG1 or PE-IgG3 stained cells within the GFP^+ population. **(b)** Compilation of CSR results in *aid-/-* and *aid-/- msh2-/-* cells expressing RV-AID-ER, RV-ΔAID-ER, or RV-ER showing the mean+SEM for 9 cultures, 3 cultures per mouse. **(c)** FACS results showing CSR to IgG1 and IgG3 (indicated on left) in *aid-/-* and *aid-/-msh6-/-* splenic B cells expressing AID-ER, ΔAID -ER, or ER. Similar to a. **(d)** Compilation of CSR results in *aid-/-* and *aid-/- msh6-/-* cells expressing AID-ER, ΔAID-ER, or ER showing the mean+SEM for 9 cultures, 3 cultures per mouse.

The dominant negative effect of ΔAID-ER depends upon the presence of MMR proteins.

Patients heterozygous for C terminal deleted AID have hyper-IgM syndrome (HIGM2), suggesting that this mutant is a dominant negative mutant, and consistent with the known ability of the C terminal deleted AID to dimerize with full-length AID (36). We asked whether ΔAID-ER has a dominant negative effect when expressed in mouse splenic B cells with normal endogenous AID levels. We compared IgG1 and IgG3 as a function of GFP mean fluorescence intensity (MFI), as an indicator of AID expression levels. Increasing expression of AID-ER results in increased IgG1 CSR relative to cells expressing endogenous AID levels (GFP-negative) (Fig 10a, left panel). However, increasing doses of ΔAID-ER does not increase CSR to IgG1, and at the highest expression level results in decreased CSR. CSR to IgG3 does not increase significantly with increasing expression of AID-ER in WT B cells, but decreases dramatically with increasing expression of ΔAID-ER (Fig 10a, middle panel). These data demonstrate that ΔAID-ER has a dominant negative effect on CSR in WT mouse B cells.

We considered two possibilities to explain the dominant negative effect of ΔAID. Its ability to dimerize with endogenous AID might sequester endogenous AID from Sμ. Alternatively, the heterodimer might bind Sμ but prevent association of

UNG and/or MMR proteins with full-length AID at Sμ. Either mechanism would inhibit CSR. To examine these possibilities, we performed ChIP using anti-ER antibody to determine if ΔAID-ER associates with Sμ in cells expressing endogenous AID. WT B cells transduced with ER were used as negative controls. The ChIP results with anti-ER antibody indicate that ΔAID-ER is actually recruited better than AID-ER to the Sμ region in AID-sufficient cells (p=0.005)(Fig 10a, right panel).

We hypothesized that the dominant negative effect observed in WT B cells expressing ΔAID-ER might be due to the inability of heterodimers of ΔAID-ER and endogenous WT AID to recruit CSR-specific proteins, such as UNG and/or Msh2-Msh6. If this is correct, then the dominant negative effect should be reduced in MMR-deficient B cells relative to that in WT B cells. To test this, we expressed AID and Δ AID in AID-sufficient *msh2-/-* or *mlh1-/-* B cells and analyzed CSR by flow cytometry. Expression of AID-ER in WT B cells increases switching to IgG1 and IgG3 relative to cells expressing the negative control ER (Fig 10b-e). However, CSR is significantly lower in WT cells expressing ΔAID-ER than in cells expressing ER (Fig 10c, e). This illustrates the dominant negative effect of ΔAID-ER in WT B cells. Interestingly, ΔAID-ER does not have a dominant negative effect in *msh2-/-* or *mlh1-/-* cells. These results taken together with the ChIP results suggest that deletion of the C-terminus of AID

reduces the ability of endogenous AID to recruit MMR proteins to Sμ, thereby preventing the contribution of MMR to CSR.

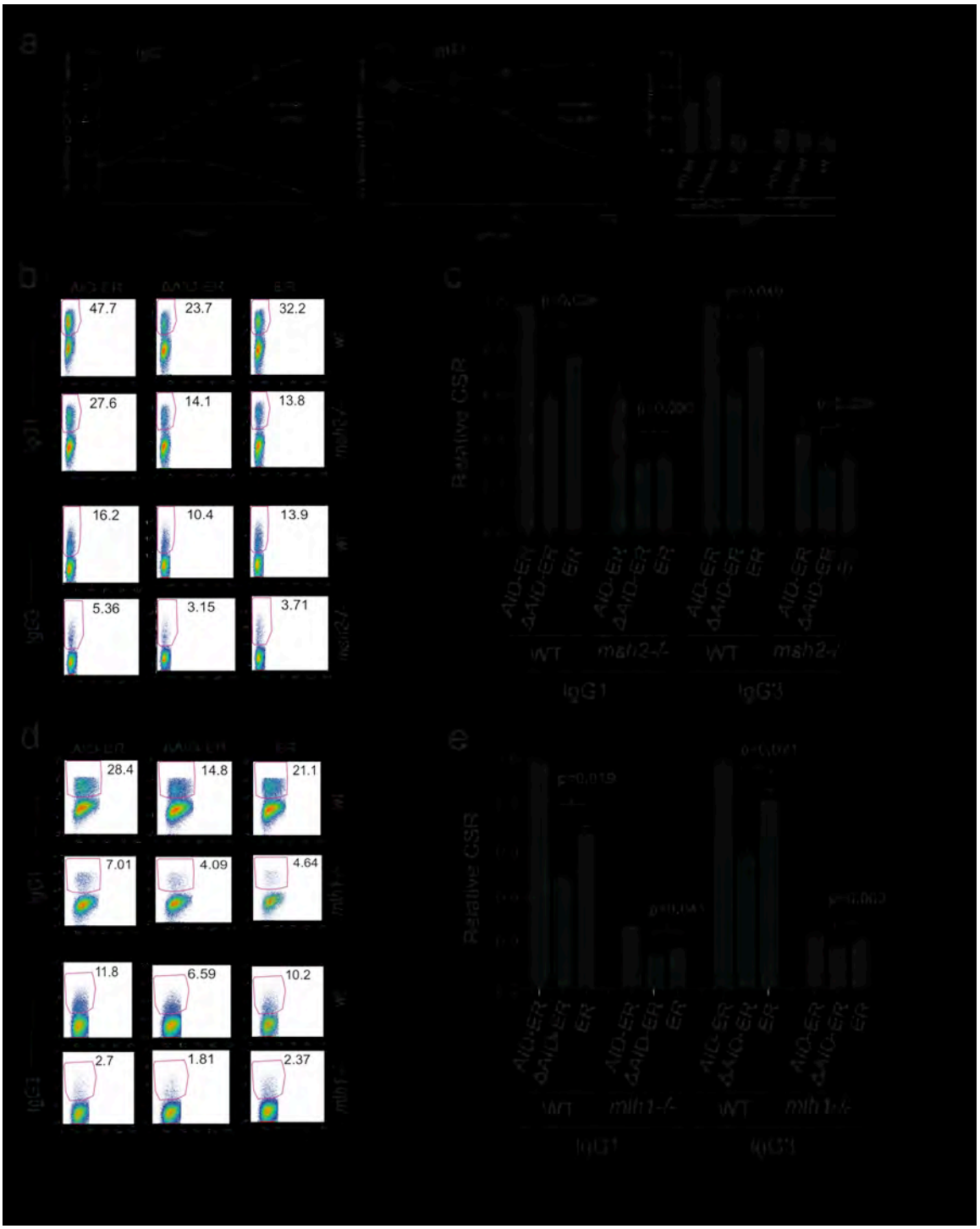
47.7
23.7
32.2
27.6
14.1
13.8
16.2
10.4
13.9
5.36
3.15
3.71
28.4
14.8
21.1
7.01
4.09
4.64
11.8
6.59
10.2
2.7
1.81
2.37
IgG1
IgG3
WT
msh2-/-
msh1-/-

Figure 10

Figure 10 ΔAID-ER inhibits recruitment of MMR proteins to Sμ by endogenous AID, thereby resulting in a dominant negative effect on CSR. **(a)** IgG1 and IgG3 switching in *aid+/+* mouse splenic B cells expressing increasing doses of AID-ER or ΔAID-ER, as assayed by GFP expression (fluorescence intensity), are increased or inhibited, respectively. Cells expressing more GFP express more AID or ΔAID (data not shown). Each data point on the curve is the mean %IgG1 or %IgG3±SEM of 3 cultures. Right panel shows an anti-ER ChIP assay of WT splenic B cells expressing AID-ER, ΔAID-ER, or ER. (**b)** FACS results for one representative CSR experiment comparing the dominant negative effect of ΔAID-ER in WT and *msh2-/-* B cells (expressing endogenous AID). The FACS plots show only viable and GFP positive cells, the gates represent PE-IgG1/PE-IgG3 positive cells within the GFP^+ populations as indicated. **(c)** Compilation of CSR results(represented in b), CSR is relative to AID-ER expressing WT B cells for 6 cultures, 3 cultures per mouse. Error bars represent SEMs. **(d)** FACS results for one experiment comparing the dominant negative effect of ΔAID-ER in WT and *mlh1-/-* B cells. **(e)** Compilation of CSR assays(represented in d), CSR is relative to AID-ER expressing WT B cells for 6 cultures, 3 cultures per mouse. Error bars represent SEMs. Lower panel: percent reduction in CSR in RV-ΔAID-ER expressing cells relative to RV-ER (dominant negative effect) in WT vs *mlh1-/-* splenic B cells for IgG1 and IgG3 CSR.

AID and UNG do not interact *in vitro*.

In order to test whether AID and UNG interact *in vitro*, I performed GST pull-down with N-terminally GST tagged UNG. I used an *in vitro* transcription/translation kit to prepare AID-ER, ΔAID-ER, ER, GST, and GST-UNG (N-terminal tagged) *in vitro*. I performed numerous control experiments, which indicated very efficient production of the above-mentioned proteins. The GST-UNG and GST alone pull-down assay yielded the same results, showing no preferential binding of full length AID-ER relative to ΔAID-ER or ER. Control experiments demonstrated that both GST and GST-UNG were pulled down efficiently by glutathione beads (fig 11b,d). I conclude that AID and UNG do not directly interact *in vitro* (fig 11a,c).

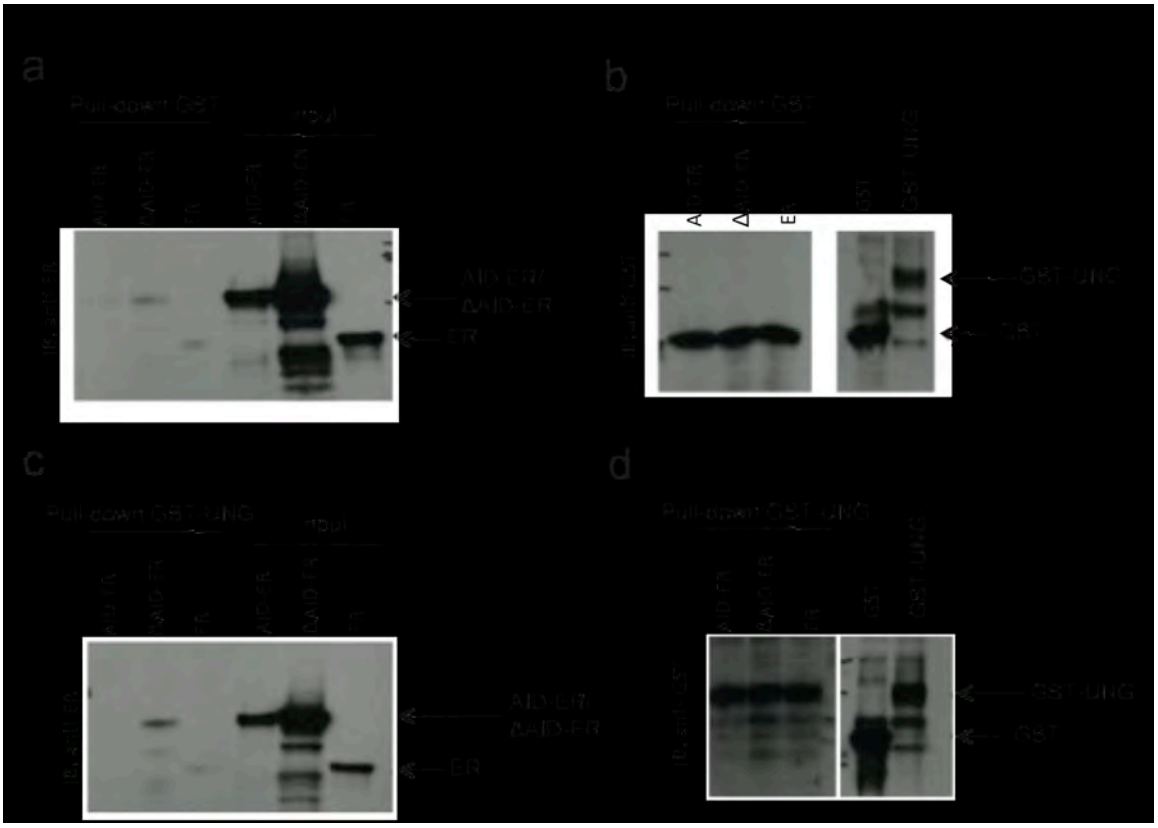

Figure 11

Figure 11 In vitro translated AID and UNG do not interact *in vitro*. (a) Immunoblot probed with anti-ER antibody shows that GST is unable to pull down in vitro translated AID-ER, ΔAID-ER or ER. **(b)** Immunoblot of aliquots from GST pull-down probed with anti-GST antibody shows efficient pull down of GST by glutathione beads. **(c)** Immunoblot probed with anti-ER antibody shows that GST-UNG is unable to pull down AID-ER, ΔAID -ER or ER *in vitro*. **(d)** Immunoblot of aliquots from GST-UNG pull-down probed with anti-GST antibody shows efficient pull down of GST-UNG by glutathione beads.

ΔAID-ER induces S region DSBs in *aid-/-* B cells.

DSBs are an essential intermediate in CSR, and are dependent upon AID, UNG (42), and MMR (12). Our findings that ΔAID-ER binds less well to Sμ than does AID-ER, and that ΔAID-ER does not recruit UNG or MMR suggest that there might be fewer Sμ DSBs in vivo in cells undergoing CSR. To address this hypothesis, we performed ligation-mediated (LM)-PCR to analyze Sμ DSBs in *aid-/-* B cells expressing AID-ER, ΔAID-ER, or ER. As expected, uninfected WT B cells showed abundant DSBs in Sμ and Sγ3 regions while *aid-/-* B cells, uninfected or infected with control RV-ER showed very few S region DSBs (Fig 12a). We also performed LM-PCR on GFP-negative cells FACS-sorted from RV-infected populations and, as expected, found very few S region DSBs (Fig 12a). Surprisingly, *aid-/-* cells expressing either AID-ER or ΔAID-ER contain abundant S region DSBs. We quantified the DSBs using densitometry, which revealed that ΔAID-ER induced more DSBs (~1.5 fold) than AID-ER in both Sμ and Sγ3 regions of *aid-/-* B cells (Fig 12b). As shown in Fig 12a, the GFP expression among the different cell populations used for LM-PCR was similar, suggesting that the expression of AID and ΔAID was similar in these cultures. The finding that DSBs are at least as numerous in ΔAID–ER- relative to AID-ER-expressing cells is surprising because both UNG and MMR are important for DSBs in S regions, and these proteins do not appear to be recruited in cells expressing ΔAID–ER. These data clearly indicate that the block in CSR due to lack of the AID C terminus occurs downstream of S region DSB formation.

To determine whether there is any difference between the DSB sites in cells expressing full length or C terminal deleted AID, we cloned and sequenced the LM-PCR products. The sites of DSBs in Sμ of AID-ER or ΔAID-ER expressing *aid-/-* B cells are indistinguishable from each other and from those in WT B cells expressing endogenous AID (10, 42) (Fig 12c). Most of the DSBs occur at G:C bp, and preferentially at AID (WRC/GYW) target hotspots. These data suggest that deletion of the AID C terminus does not alter the mechanism or the ability of AID to generate DSBs in the Sμ region. Thus, although the weak binding of ΔAID-ER relative to AID-ER to Sμ is sufficient to induce normal frequencies of deamination and DSBs in S regions, it appears insufficient to direct the S region DSBs toward proper CSR.

In conclusion, these data reveal that the C-terminus of AID is involved in recruitment of UNG and Msh2-Msh6 to Ig Sμ regions, and that this recruitment is important for the function of Msh2-Msh6 in CSR. Although both UNG and MMR have been demonstrated to be important for DSB formation during CSR, our results indicate that the AID C terminus is not required for Sμ or Sγ3 DSBs. Our results support the hypothesis that in addition to deamination of dCs, leading to DSBs, AID has functions downstream of generation of DSBs that are required for CSR, and that this involves specific recruitment of UNG and MMR to S regions.

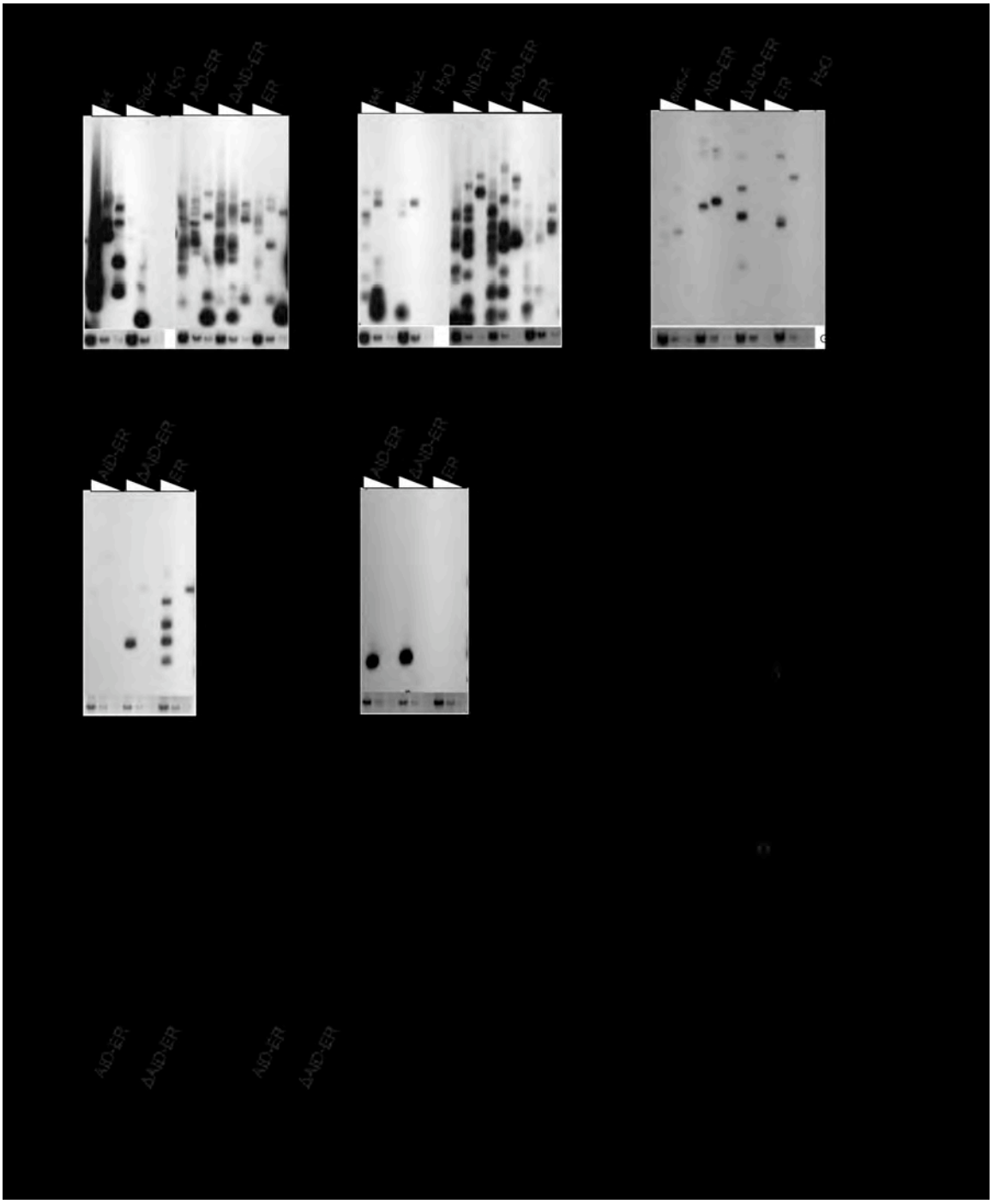

Figure 12

Figure 12 ΔAID-ER is proficient in inducing Sμ and Sγ3 DSBs during CSR. **(a)** Upper panels: Southern blots showing LM-PCR results for *aid-/-* splenic B cells expressing AID-ER, ΔAID-ER, or ER cultured to switch to IgG3. LM-PCRs performed on DNA from WT and *aid-/-* B cells were used for positive and negative controls, respectively. Lower panels show LM-PCR results for GFP-negative cells from the same *aid-/-* cultures shown in the upper panels, demonstrating that DSBs depend on AID and ΔAID. GAPDH indicates the PCR for *gapdh* genes, used as an internal template loading control. 3-fold dose titrations were assayed. Mean GFP expression levels+SEM for AID-ER, ΔAID-ER and ER cultures in 3 independent experiments (different mice) are shown in the histograms. **(b)** Quantitation of 3 LM-PCR experiments by densitometry of autoradiographic films, scanning all 3 lanes for each genotype; histograms represent relative means+SEM. **(c)** Tabulation of sites of blunt DSBs obtained from LM-PCR products, comparing results from cells expressing endogenous AID (10).

Chapter IV

DISCUSSION

It is intriguing that deamination of S-regions by AID is not the entire role of AID during CSR. Although ΔAID appears to induce normal levels of deamination in the S-regions when B cells are activated to switch, the deamination is not translated into a normal levels of CSR (34). This phenomenon makes it quite clear that AID has a larger role during CSR than simply deamination. This prompted me to study the role of AID in steps of CSR that are downstream of deamination.

Previous studies (observation in humans) (36) and my observations in this study (in mouse) have proved that ΔAID inhibits CSR when expressed in WT B cells. The dominant negative effect of ΔAID is a clear indication that ΔAID interferes with full length AID's ability to support CSR, which I speculate to require interactions between AID and other proteins required for CSR. The interaction between AID and other CSR-specific proteins probably not only promotes efficient completion of CSR, but also channels the AID protein away from its potent genotoxic property, for example denying the catalytic site of AID to access the DNA by steric hinderance. AID is an essential protein for the humoral immune response, yet it's activity is highly genotoxic. Hence, AID protein and B

cells must have evolved to balance the necessary activity and genotoxic activity of AID. In support of this hypothesis, it is clear from previous studies that B cells have evolved to allow high levels of AID expression only upon immunization. B cells have evolved to compartmentalize AID protein in the cytoplasm by making the nuclear localization signal weak (32), and nuclear export signal strong (39). They have also evolved to make sure that the deamination activity is potentiated by post-translation modification such as phosphorylation of S38; only a small percentage of nuclear AID is phosphorylated (72-74). In addition, cells have evolved to rapidly ubiquitinate and degrade nuclear AID; nuclear AID is three times less stable than cytoplasmic AID (76).

It is clear from my data that a transient interaction (undetectable by ChIP) of AID with DNA is adequate to generate mutations at the S-region; hence, any longer interaction of AID protein with DNA would be highly detrimental. Hence, keeping in mind the potent deamination activity of AID and the need for the AID activity to be negatively regulated, I thereby speculate that AID has evolved to interact with other proteins that are necessary for CSR to not only dampen the deamination activity of AID, but also to localize the deletional recombination to the *Igh* locus. Deletional recombination is essential for CSR but it is highly detrimental to the cells if it takes place between *Igh* and another locus or among other loci. AID induces significant collateral damage during CSR, as evidenced by AID induced DNA breaks in other loci (100-103). However, to ensure that the deletional

recombination among those loci doesn't take place, cells might have evolved to employ AID itself to form a CSR specific recombination complex. Hence, AID might have evolved to nucleate a CSR specific complex by co-dependent binding of AID with other proteins necessary for CSR at the *Igh* locus.

To test this hypothesis, I immunoprecipitated (IP) AID-ER protein, expressed in B cells with the help of retroviral transduction, with anti-ER antibody and tested for co-IP of proteins that are necessary for CSR such as UNG and Msh2/Msh6. The IP however did not yield any positive co-IP. Unsatisfied with this revelation, I decided to find out whether AID protein intrinsically interacts with UNG protein. For this end, I utilized *in vitro* transcribed and translated AID-ER, ΔAID-ER, ER, GST and GST-UNG (N-terminal GST) proteins; the ITT yielded copious amounts of the proteins. The GST pull down of the UNG protein revealed that AID and UNG do not have intrinsic ability to interact with each other.

Since I was unable to detect protein-protein interactions among AID and other CSR specific proteins by IP, I decided to dissect the protein-DNA-protein interaction by ChIP assay. In this study, I have demonstrated that AID binds co-dependently with both UNG and Msh2-Msh6 to the Ig Sμ region in B cells, and that the co-dependent binding requires the C terminal 10 amino acids of AID. No binding of AID, UNG or Msh2-Msh6 was detected at the Cμ gene. The associations between AID and UNG and between AID and Msh2-Msh6 are

detected when AID is localized on the Sμ region; I was not able to detect binding in the absence of DNA. The interaction might be direct or indirect. Interestingly, we found that the MMR proteins Msh2 and Mlh1 do not contribute to CSR in cells expressing C terminal deleted AID, indicating that the ability of AID to recruit MMR proteins to Sμ is important for CSR. We also showed that ΔAID is a dominant negative mutant, most likely due to its ability to dimerize with full-length AID and to prevent recruitment of Msh2-Msh6, and perhaps other proteins, to the Sμ region. The C terminal deletion mutant might also prevent recruitment of UNG by endogenous full-length AID, but since CSR is nearly abolished in *ung-/-* cells (6, 12), we were unable to monitor this dominant negative effect functionally.

It is puzzling that there are at least as many Sμ DSBs, if not more, in cells expressing ΔAID as in cells expressing full-length AID, despite lower recruitment of AID, UNG and MMR. Sμ region DSBs have also been observed in HIGM patients with C terminal deletions of AID (104). Considerable evidence indicates that MMR is important for conversion of SSBs to DSBs in S regions during CSR (11, 12, 105). In its absence DSBs appear to only form from SSBs that are near on opposite strands. The observed DSBs in ΔAID-expressing cells might be due to such nearby SSBs. I hypothesize that the slightly increased Sμ DSBs and ablation of CSR in cells expressing C terminal deleted AID are due to impaired downstream DNA repair and/or recombination. Unresolved or inefficient joining

of DSBs could cause accumulation of S region DSBs, and also could cause increased risks of DNA translocations, carcinogenesis, and/or cell death. In fact, ΔAID-ER-expressing B cells have a higher frequency of chromosomal translocations than cells expressing AID-ER (106).

Chromosomal translocations have been shown to be mediated by microhomology-mediated end joining (107), a non-classical form of NHEJ in which recombination is mediated by short microhomologies between the two recombining sequences. In addition, B cells with mutations in DNA repair genes involved in recombination of S region breaks often show increased lengths of junctional microhomology (108). Deficiencies in some MMR proteins also increase junctional microhomology (108-110). Interestingly, it was recently found that Sμ-Sα junctions from two human patients expressing AID with C terminal deletions show highly significantly increased junctional microhomology (111). These results suggest that the AID C terminus is involved in the recombination step during CSR and directs S region DSBs towards NHEJ.

Interaction of AID with UNG and MMR might direct S region DSBs towards properly regulated CSR. AID-dependent Sμ DSBs are only observed during G1 phase in normal splenic B cells undergoing CSR, indicating that the DBS are introduced, recombined and repaired during G1 phase (12, 43). MMR associates with the replication machinery and corrects nucleotides mis-

incorporated due to DNA polymerase errors (13, 112). UNG is important for removal of dU incorporated instead of dT during replication, and in both primary fibroblasts and HeLa cells UNG is expressed at much higher levels during S phase than during G1 or G2 phases (113, 114). However, in splenic B cells induced to undergo CSR, UNG is slightly more highly expressed in G1 phase than in S phase, and is clearly important for excision of dUs due to deamination by AID (12, 42). Perhaps UNG is stabilized by binding to AID. It is possible that if UNG and MMR are not specifically recruited to Sμ during G1 phase, they might only be available at S regions during S phase, thus creating DSBs in S phase. During DNA replication, DSBs cause stalling of DNA polymerase and collapse of replication forks, which could result in unrepaired DSBs, increased translocations, and greatly reduced CSR.

The binding of UNG and MMR to AID might be indirect and occur via other proteins. AID has been shown to interact with several proteins. Protein kinase A (PKA), which phosphorylates AID at S38, is specifically localized at Sμ regions during CSR, and might participate in a PKA-RPA-AID complex (2, 85). Phosphorylation of AID at S38 is required for interaction with RPA and for CSR (71, 86), but it is possible that the AID C terminus is also involved. Interestingly, RPA32 has been shown to interact with UNG by the yeast two-hybrid assay (88). Also, a recent publication reported that the C-terminus of AID interacts with the 14-3-3 adaptor protein complex and that this interaction is important for

recruitment of AID to S regions (78). Deficiency in 14-3-3 protein components results in a 50% decrease in CSR. Since the ΔAID-ER mutant induces as many Sμ mutations as AID-ER, the recruitment by 14-3-3 of AID itself is not necessary for the ability of AID to deaminate Sμ. It is possible, however, that 14-3-3, when bound to AID, is important for recruiting UNG and MMR proteins to Sμ.

UNG and MMR proteins might help recruit other proteins involved in subsequent repair steps during CSR. For example, Msh2 and Msh6 have been found in a large complex of DNA repair proteins in HeLa cells, termed BASC, which also includes Mre11-Rad50-Nbs1 (MRN). MRN binds DSBs induced by AID and is known to be important for CSR (1, 115, 116). UNG has been found in a complex in HeLa cells that contains proteins involved in base excision repair, including APE1, XRCC1, DNA polymerase β, PCNA, and DNA ligase (117). APE1 is important for CSR, and is involved in converting AID-induced lesions to DNA breaks (10). DNA Pol β binds S regions, decreases the numbers of Sμ DSBs, and inhibits CSR (118). However, it is difficult to understand how these proteins would direct the DSBs toward proper S-S recombination. Another possibility is suggested by the findings in over-expression experiments that AID binds DNA-PKcs, dependent upon the AID C terminus (79), and that Msh2-Msh6 binds to Ku70-Ku80 (119). As Ku is important for directing DNA breaks towards NHEJ, it is possible that in addition to its role in converting SSBs to DSBs, MMR

associated with the AID C terminus is involved in recruiting DNA-PK-Ku proteins to sites of AID-induced DSBs.

Here I propose a model for the initiation of CSR that incorporates my new results (Fig 13). After immunization, transcription through S-region occurs to generate ssDNA substrate for AID. AID is induced and a minimal amount is translocated into the nucleus, a portion of which is phosphorylated at S38, which allows AID to interact with RPA32, allowing it to bind more strongly to DNA and making its deamination activity more potent. The phosphorylated form has been shown to be located preferentially in the chromatin fraction (72-74). Meanwhile AID interacts with 14-3-3 adaptor proteins, which altogether promote recruitment of UNG and MMR protein to the S regions, to form a protein complex. Additional DNA repair proteins might be recruited. I propose that this complex is responsible for further steps required for CSR, eventually resulting in regulated and proper S-S recombination.

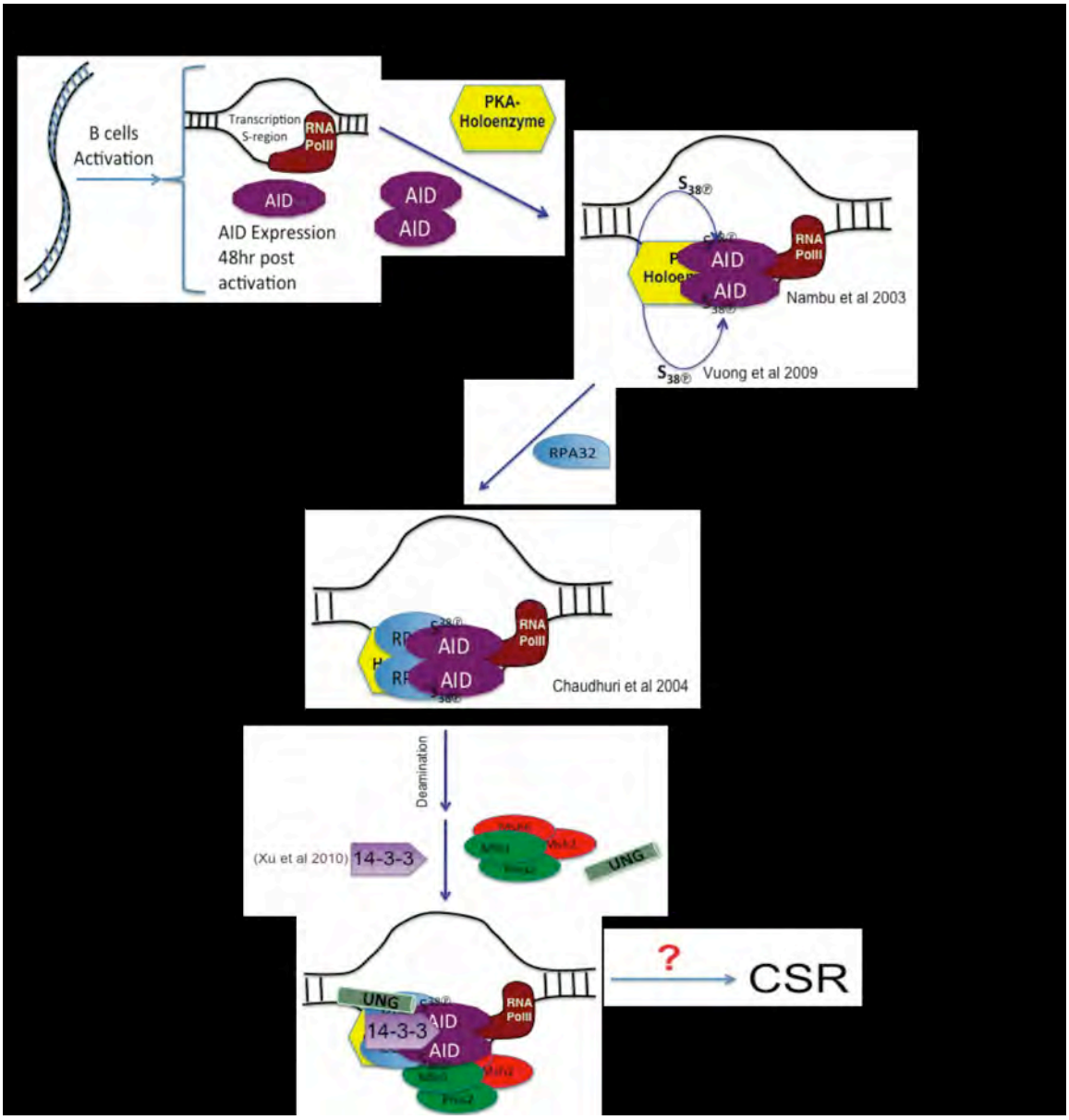
B cells
Activation
Transcription
S-region
RNA
PolII
AID
AID Expression
48hr post
activation
PKA-
Holoenzyme
AID
AID
S38
AID
AID
RNA
PolII
Nambu et al 2003
S38
Vuong et al 2009
RPA32
AID
AID
RNA
PolII
Chaudhuri et al 2004
Deamination
(Xu et al 2010)
14-3-3
UNG
UNG
14-3-3
AID
AID
RNA
PolII
?
CSR

Figure 13

Figure 13 Model for CSR complex assembly. After the activation of B cells via immunization of infection, transcription through donor and acceptor S-regions take place and AID is expressed. AID and (DNA dependent protein kinase A) PKA are recruited on the S-region independent of each other. PKA then phosphorylates AID in S38, which promotes AID interaction with RPA32. AID-RPA32 interaction DNA binding as well as catalytic activity of AID. This complex is deamination competent. Then AID interacts with 14-3-3 adaptor protein in AID-C-terminus dependent manner, which stabilizes AID-DNA complex. Then UNG and MMR proteins are also recruited into this complex in the AID-C-terminus dependent manner as well. This super complex is not needed for DSB during CSR however, it is essential for one or more steps that lies downstream of DSB formation step.

Future directions

To further solidify the idea of co-dependent binding, one could perform some more control experiments such as ChIP with anti-UNG antibody in MMR deficient B cells and vice versa. It is also important to test whether other DNA repair factors demonstrate AID-dependent co-dependent binding with S-regions. This can be accomplished by ChIP as was done in this study. This will require generating various DNA repair and AID double deficient mice.

Although it is clear that co-dependent binding of AID-UNG-MMR to the Sμ region is important for CSR, it is also clear that this co-dependent interaction among the proteins is not essential for the deamination step nor the DSB formation step of CSR. Hence, it is important to dissect the steps downstream of both deamination and DSBs during CSR. In order to address whether the NHEJ/DNA repair step of CSR is compromised in ΔAID expressing cells, one can try to rescue CSR in ΔAID expressing cells by, for example, homologous recombination. This can be accomplished by cloning homologous recombination cassettes containing AID hotspot motifs (WRC) into a donor and acceptor S region and testing whether the CSR can be rescued by the cassettes in ΔAID expressing cells.

In order to test whether a homologous recombination cassette can rescue CSR in ΔAID expressing *aid-/-* mouse B cells, a knock-in (KI) mouse can be created in

which Sμ and Sγ3 regions are replaced by homologous recombination cassettes. These cassettes should contain AID hotspots so that AID can introduce mutations to initiate DSB formation process. The KI-mice can be then crossed with *aid-/-* mice to generate KI X *aid*-/- mice. Then AID-ER, ΔAID-ER or ER can be expressed in these KI X *aid-/-* mouse splenic B cells with the help of retroviral transduction. The phenotype of the transduced B cells can then be evaluated by CSR assay as we have done in this study. If the presence of homologous recombination cassettes in the B cells rescue CSR in ΔAID-ER expressing *aid-/-* B cells, then it would mean that AID protein is important for carrying out recombination step of CSR but Δ AID is incompetent in carrying out the recombination step. If the homologous recombination cassette doesn't rescue the CSR, then it could mean that ΔAID is probably incompetent in the CSR step that is upstream of recombination step, which is not yet characterized.

It is not clear what roles UNG and MMR proteins might have during the CSR steps downstream of the DSB formation. It is possible that the interaction of these proteins with AID might furnish these proteins with non-traditional roles in, for example, locus specific recombination. Hence, it is important to develop assays that can dissect the roles of UNG and MMR proteins in the CSR steps downstream of DSB formation. And it would be very important to develop another assay to understand the biochemical basis of their roles in these processes.

REFERENCES

1. **Stavnezer, J., J.E. Guikema, and C.E. Schrader. 2008. Mechanism and regulation of class switch recombination. *Annu Rev Immunol* 26:261-292.**
2. **Chaudhuri, J., C. Khuong, and F.W. Alt. 2004. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* 430:992-998.**
3. **Dickerson, S.K., E. Market, E. Besmer, and F.N. Papavasiliou. 2003. AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med* 197:1291-1296.**
4. **Pham, P., R. Bransteitter, J. Petruska, and M.F. Goodman. 2003. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* 424:103-107.**
5. **Bransteitter, R., P. Pham, M.D. Scharff, and M.F. Goodman. 2003. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci U S A* 100:4102-4107.**
6. **Rada, C., G.T. Williams, H. Nilsen, D.E. Barnes, T. Lindahl, and M.S. Neuberger. 2002. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr Biol* 12:1748-1755.**
7. **Imai, K., G. Slupphaug, W.I. Lee, P. Revy, S. Nonoyama, N. Catalan, L. Yel, M. Forveille, B. Kavli, H.E. Krokan, H.D. Ochs, A. Fischer, and A. Durandy. 2003. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. *Nat Immunol* 4:1023-1028.**
8. **Di Noia, J., and M.S. Neuberger. 2002. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* 419:43-48.**
9. **Ramiro, A.R., M. Jankovic, E. Callen, S. Difilippantonio, H.T. Chen, K.M. McBride, T.R. Eisenreich, J. Chen, R.A. Dickins, S.W. Lowe, A. Nussenzweig, and M.C. Nussenzweig. 2006. Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. *Nature* 440:105-109.**
10. **Guikema, J.E., E.K. Linehan, D. Tsuchimoto, Y. Nakabeppu, P.R. Strauss, J. Stavnezer, and C.E. Schrader. 2007. APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination. *J Exp Med* 204:3017-3026.**
11. **Stavnezer, J., and C.E. Schrader. 2006. Mismatch repair converts AID-instigated nicks to double-strand breaks for antibody class-switch recombination. *Trends Genet* 22:23-28.**

12. **Schrader, C.E., J.E. Guikema, E.K. Linehan, E. Selsing, and J. Stavnezer. 2007. Activation-induced cytidine deaminase-dependent DNA breaks in class switch recombination occur during G1 phase of the cell cycle and depend upon mismatch repair. *J Immunol* 179:6064-6071.**
13. **Modrich, P. 2006. Mechanisms in eukaryotic mismatch repair. *J Biol Chem* 281:30305-30309.**
14. **Kunkel, T.A., and D.A. Erie. 2005. DNA mismatch repair. *Annu Rev Biochem* 74:681-710.**
15. **Bardwell, P.D., C.J. Woo, K. Wei, Z. Li, A. Martin, S.Z. Sack, T. Parris, W. Edelmann, and M.D. Scharff. 2004. Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. *Nat Immunol* 5:224-229.**
16. **van Oers, J.M., S. Roa, U. Werling, Y. Liu, J. Genschel, H. Hou, Jr., R.S. Sellers, P. Modrich, M.D. Scharff, and W. Edelmann. PMS2 endonuclease activity has distinct biological functions and is essential for genome maintenance. *Proc Natl Acad Sci U S A* 107:13384-13389.**
17. **Kadyrov, F.A., L. Dzantiev, N. Constantin, and P. Modrich. 2006. Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* 126:297-308.**
18. **Petersen-Mahrt, S.K., R.S. Harris, and M.S. Neuberger. 2002. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* 418:99-103.**
19. **Zeng, X., D.B. Winter, C. Kasmer, K.H. Kraemer, A.R. Lehmann, and P.J. Gearhart. 2001. DNA polymerase eta is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat Immunol* 2:537-541.**
20. **Barreto, V.M., A.R. Ramiro, and M.C. Nussenzweig. 2005. Activation-induced deaminase: controversies and open questions. *Trends Immunol* 26:90-96.**
21. **Li, Z., C.J. Woo, M.D. Iglesias-Ussel, D. Ronai, and M.D. Scharff. 2004. The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev* 18:1-11.**
22. **Maizels, N. 2005. Immunoglobulin gene diversification. *Annu Rev Genet* 39:23-46.**
23. **Durandy, A., P. Revy, and A. Fischer. 2004. Human models of inherited immunoglobulin class switch recombination and somatic hypermutation defects (hyper-IgM syndromes). *Adv Immunol* 82:295-330.**
24. **Martin, A., Z. Li, D.P. Lin, P.D. Bardwell, M.D. Iglesias-Ussel, W. Edelmann, and M.D. Scharff. 2003. Msh2 ATPase activity is essential for somatic hypermutation at a-T basepairs and for efficient class switch recombination. *J Exp Med* 198:1171-1178.**

25. Frey, S., B. Bertocci, F. Delbos, L. Quint, J.C. Weill, and C.A. Reynaud. 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.

26. Rada, C., M.R. Ehrenstein, M.S. Neuberger, and C. Milstein. 1998. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity* 9:135-141.

27. Li, Z., S.J. Scherer, D. Ronai, M.D. Iglesias-Ussel, J.U. Peled, P.D. Bardwell, M. Zhuang, K. Lee, A. Martin, W. Edelmann, and M.D. Scharff. 2004. Examination of Msh6- and Msh3-deficient mice in class switching reveals overlapping and distinct roles of MutS homologues in antibody diversification. *J Exp Med* 200:47-59.

28. Martomo, S.A., W.W. Yang, and P.J. Gearhart. 2004. A role for Msh6 but not Msh3 in somatic hypermutation and class switch recombination. *J Exp Med* 200:61-68.

29. Wiesendanger, M., B. Kneitz, W. Edelmann, and M.D. Scharff. 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.

30. Li, Z., C. Zhao, M.D. Iglesias-Ussel, Z. Polonskaya, M. Zhuang, G. Yang, Z. Luo, W. Edelmann, and M.D. Scharff. 2006. The mismatch repair protein Msh6 influences the in vivo AID targeting to the Ig locus. *Immunity* 24:393-403.

31. Muto, T., M. Muramatsu, M. Taniwaki, K. Kinoshita, and T. Honjo. 2000. Isolation, tissue distribution, and chromosomal localization of the human activation-induced cytidine deaminase (AID) gene. *Genomics* 68:85-88.

32. Ito, S., H. Nagaoka, R. Shinkura, N. Begum, M. Muramatsu, M. Nakata, and T. Honjo. 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.

33. Shinkura, R., S. Ito, N.A. Begum, H. Nagaoka, M. Muramatsu, K. Kinoshita, Y. Sakakibara, H. Hijikata, and T. Honjo. 2004. Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat Immunol* 5:707-712.

34. Barreto, V., B. Reina-San-Martin, A.R. Ramiro, K.M. McBride, and M.C. Nussenzweig. 2003. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol Cell* 12:501-508.

35. Revy, P., T. Muto, Y. Levy, F. Geissmann, A. Plebani, O. Sanal, N. Catalan, M. Forveille, R. Dufourcq-Labelouse, A. Gennery, I. Tezcan, F. Ersoy, H. Kayserili, A.G. Ugazio, N. Brousse, M. Muramatsu, L.D. Notarangelo, K. Kinoshita, T. Honjo, A. Fischer, and A. Durandy.

2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* 102:565-575.

36. Ta, V.T., H. Nagaoka, N. Catalan, A. Durandy, A. Fischer, K. Imai, S. Nonoyama, J. Tashiro, M. Ikegawa, S. Ito, K. Kinoshita, M. Muramatsu, and T. Honjo. 2003. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat Immunol* 4:843-848.

37. Zhu, Y., S. Nonoyama, T. Morio, M. Muramatsu, T. Honjo, and S. Mizutani. 2003. Type two hyper-IgM syndrome caused by mutation in activation-induced cytidine deaminase. *J Med Dent Sci* 50:41-46.

38. Brar, S.S., M. Watson, and M. Diaz. 2004. Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. *J Biol Chem* 279:26395-26401.

39. McBride, K.M., V. Barreto, A.R. Ramiro, P. Stavropoulos, and M.C. Nussenzweig. 2004. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J Exp Med* 199:1235-1244.

40. Chaudhuri, J., M. Tian, C. Khuong, K. Chua, E. Pinaud, and F.W. Alt. 2003. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* 422:726-730.

41. Sohail, A., J. Klapacz, M. Samaranayake, A. Ullah, and A.S. Bhagwat. 2003. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Res* 31:2990-2994.

42. Schrader, C.E., E.K. Linehan, S.N. Mochegova, R.T. Woodland, and J. Stavnezer. 2005. Inducible DNA breaks in Ig S regions are dependent on AID and UNG. *J Exp Med* 202:561-568.

43. Petersen, S., R. Casellas, B. Reina-San-Martin, H.T. Chen, M.J. Difilippantonio, P.C. Wilson, L. Hanitsch, A. Celeste, M. Muramatsu, D.R. Pilch, C. Redon, T. Ried, W.M. Bonner, T. Honjo, M.C. Nussenzweig, and A. Nussenzweig. 2001. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature* 414:660-665.

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

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

46. Ramiro, A.R., P. Stavropoulos, M. Jankovic, and M.C. Nussenzweig. 2003. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat Immunol* 4:452-456.

47. Nambu, Y., M. Sugai, H. Gonda, C.G. Lee, T. Katakai, Y. Agata, Y. Yokota, and A. Shimizu. 2003. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science* 302:2137-2140.

48. Wang, L., R. Wuerffel, S. Feldman, A.A. Khamlichi, and A.L. Kenter. 2009. S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. *J Exp Med* 206:1817-1830.

49. Woo, C.J., A. Martin, and M.D. Scharff. 2003. Induction of somatic hypermutation is associated with modifications in immunoglobulin variable region chromatin. *Immunity* 19:479-489.

50. Larijani, M., A.P. Petrov, O. Kolenchenko, M. Berru, S.N. Krylov, and A. Martin. 2007. AID associates with single-stranded DNA with high affinity and a long complex half-life in a sequence-independent manner. *Mol Cell Biol* 27:20-30.

51. Imai, K., Y. Zhu, P. Revy, T. Morio, S. Mizutani, A. Fischer, S. Nonoyama, and A. Durandy. 2005. Analysis of class switch recombination and somatic hypermutation in patients affected with autosomal dominant hyper-IgM syndrome type 2. *Clin Immunol* 115:277-285.

52. Pauklin, S., I.V. Sernandez, G. Bachmann, A.R. Ramiro, and S.K. Petersen-Mahrt. 2009. Estrogen directly activates AID transcription and function. *J Exp Med* 206:99-111.

53. Gourzi, P., T. Leonova, and F.N. Papavasiliou. 2006. A role for activation-induced cytidine deaminase in the host response against a transforming retrovirus. *Immunity* 24:779-786.

54. Okazaki, I.M., H. Hiai, N. Kakazu, S. Yamada, M. Muramatsu, K. Kinoshita, and T. Honjo. 2003. Constitutive expression of AID leads to tumorigenesis. *J Exp Med* 197:1173-1181.

55. Rucci, F., L. Cattaneo, V. Marrella, M.G. Sacco, C. Sobacchi, F. Lucchini, S. Nicola, S. Della Bella, M.L. Villa, L. Imberti, F. Gentili, C. Montagna, C. Tiveron, L. Tatangelo, F. Facchetti, P. Vezzoni, and A. Villa. 2006. Tissue-specific sensitivity to AID expression in transgenic mouse models. *Gene* 377:150-158.

56. Muto, T., I.M. Okazaki, S. Yamada, Y. Tanaka, K. Kinoshita, M. Muramatsu, H. Nagaoka, and T. Honjo. 2006. Negative regulation of activation-induced cytidine deaminase in B cells. *Proc Natl Acad Sci U S A* 103:2752-2757.

57. Ramiro, A.R., M. Jankovic, T. Eisenreich, S. Difilippantonio, S. Chen-Kiang, M. Muramatsu, T. Honjo, A. Nussenzweig, and M.C. Nussenzweig. 2004. AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* 118:431-438.

58. Kotani, A., N. Kakazu, T. Tsuruyama, I.M. Okazaki, M. Muramatsu, K. Kinoshita, H. Nagaoka, D. Yabe, and T. Honjo. 2007. Activation-

induced cytidine deaminase (AID) promotes B cell lymphomagenesis in Emu-cmyc transgenic mice. *Proc Natl Acad Sci U S A* 104:1616-1620.

59. Yadav, A., A. Olaru, M. Saltis, A. Setren, J. Cerny, and F. Livak. 2006. Identification of a ubiquitously active promoter of the murine activation-induced cytidine deaminase (AICDA) gene. *Mol Immunol* 43:529-541.

60. Gonda, H., M. Sugai, Y. Nambu, T. Katakai, Y. Agata, K.J. Mori, Y. Yokota, and A. Shimizu. 2003. The balance between Pax5 and Id2 activities is the key to AID gene expression. *J Exp Med* 198:1427-1437.

61. Dedeoglu, F., B. Horwitz, J. Chaudhuri, F.W. Alt, and R.S. Geha. 2004. Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NFkappaB. *Int Immunol* 16:395-404.

62. Park, S.R., H. Zan, Z. Pal, J. Zhang, A. Al-Qahtani, E.J. Pone, Z. Xu, T. Mai, and P. Casali. 2009. HoxC4 binds to the promoter of the cytidine deaminase AID gene to induce AID expression, class-switch DNA recombination and somatic hypermutation. *Nat Immunol* 10:540-550.

63. Zhou, C., A. Saxon, and K. Zhang. 2003. Human activation-induced cytidine deaminase is induced by IL-4 and negatively regulated by CD45: implication of CD45 as a Janus kinase phosphatase in antibody diversification. *J Immunol* 170:1887-1893.

64. Lin, K.I., C. Angelin-Duclos, T.C. Kuo, and K. Calame. 2002. Blimp-1-dependent repression of Pax-5 is required for differentiation of B cells to immunoglobulin M-secreting plasma cells. *Mol Cell Biol* 22:4771-4780.

65. Shaffer, A.L., K.I. Lin, T.C. Kuo, X. Yu, E.M. Hurt, A. Rosenwald, J.M. Giltnane, L. Yang, H. Zhao, K. Calame, and L.M. Staudt. 2002. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity* 17:51-62.

66. Klein, U., S. Casola, G. Cattoretti, Q. Shen, M. Lia, T. Mo, T. Ludwig, K. Rajewsky, and R. Dalla-Favera. 2006. Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat Immunol* 7:773-782.

67. Sciammas, R., A.L. Shaffer, J.H. Schatz, H. Zhao, L.M. Staudt, and H. Singh. 2006. Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity* 25:225-236.

68. Teng, G., P. Hakimpour, P. Landgraf, A. Rice, T. Tuschl, R. Casellas, and F.N. Papavasiliou. 2008. MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity* 28:621-629.

69. Dorsett, Y., K.M. McBride, M. Jankovic, A. Gazumyan, T.H. Thai, D.F. Robbiani, M. Di Virgilio, B.R. San-Martin, G. Heidkamp, T.A.

Schwickert, T. Eisenreich, K. Rajewsky, and M.C. Nussenzweig. 2008. MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity* 28:630-638.

70. de Yebenes, V.G., L. Belver, D.G. Pisano, S. Gonzalez, A. Villasante, C. Croce, L. He, and A.R. Ramiro. 2008. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J Exp Med* 205:2199-2206.

71. Basu, U., J. Chaudhuri, C. Alpert, S. Dutt, S. Ranganath, G. Li, J.P. Schrum, J.P. Manis, and F.W. Alt. 2005. The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature* 438:508-511.

72. McBride, K.M., A. Gazumyan, E.M. Woo, V.M. Barreto, D.F. Robbiani, B.T. Chait, and M.C. Nussenzweig. 2006. Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. *Proc Natl Acad Sci U S A* 103:8798-8803.

73. Pasqualucci, L., Y. Kitaura, H. Gu, and R. Dalla-Favera. 2006. PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. *Proc Natl Acad Sci U S A* 103:395-400.

74. Chatterji, M., S. Unniraman, K.M. McBride, and D.G. Schatz. 2007. Role of activation-induced deaminase protein kinase A phosphorylation sites in Ig gene conversion and somatic hypermutation. *J Immunol* 179:5274-5280.

75. Rada, C., J.M. Jarvis, and C. Milstein. 2002. AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc Natl Acad Sci U S A* 99:7003-7008.

76. Aoufouchi, S., A. Faili, C. Zober, O. D'Orlando, S. Weller, J.C. Weill, and C.A. Reynaud. 2008. Proteasomal degradation restricts the nuclear lifespan of AID. *J Exp Med* 205:1357-1368.

77. Geisberger, R., C. Rada, and M.S. Neuberger. 2009. The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proc Natl Acad Sci U S A* 106:6736-6741.

78. Xu, Z., Z. Fulop, G. Wu, E.J. Pone, J. Zhang, T. Mai, L.M. Thomas, A. Al-Qahtani, C.A. White, S.R. Park, P. Steinacker, Z. Li, J. Yates, 3rd, B. Herron, M. Otto, H. Zan, H. Fu, and P. Casali. 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. *Nat Struct Mol Biol* 17:1124-1135.

79. Wu, X., P. Geraldes, J.L. Platt, and M. Cascalho. 2005. The double-edged sword of activation-induced cytidine deaminase. *J Immunol* 174:934-941.

80. Franco, S., M.M. Murphy, G. Li, T. Borjeson, C. Boboila, and F.W. Alt. 2008. DNA-PKcs and Artemis function in the end-joining phase of

immunoglobulin heavy chain class switch recombination. *J Exp Med* 205:557-564.

81. Pavri, R., A. Gazumyan, M. Jankovic, M. Di Virgilio, I. Klein, C. Ansarah-Sobrinho, W. Resch, A. Yamane, B.R. San-Martin, V. Barreto, T.J. Nieland, D.E. Root, R. Casellas, and M.C. Nussenzweig. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* 143:122-133.

82. Conticello, S.G., K. Ganesh, K. Xue, M. Lu, C. Rada, and M.S. Neuberger. 2008. Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNBL1. *Mol Cell* 31:474-484.

83. Han, L., S. Masani, and K. Yu. Cutting edge: CTNNBL1 is dispensable for Ig class switch recombination. *J Immunol* 185:1379-1381.

84. MacDuff, D.A., M.S. Neuberger, and R.S. Harris. 2006. MDM2 can interact with the C-terminus of AID but it is inessential for antibody diversification in DT40 B cells. *Mol Immunol* 43:1099-1108.

85. Vuong, B.Q., M. Lee, S. Kabir, C. Irimia, S. Macchiarulo, G.S. McKnight, and J. Chaudhuri. 2009. Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. *Nat Immunol* 10:420-426.

86. McBride, K.M., A. Gazumyan, E.M. Woo, T.A. Schwickert, B.T. Chait, and M.C. Nussenzweig. 2008. Regulation of class switch recombination and somatic mutation by AID phosphorylation. *J Exp Med* 205:2585-2594.

87. Kohli, R.M., S.R. Abrams, K.S. Gajula, R.W. Maul, P.J. Gearhart, and J.T. Stivers. 2009. A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. *J Biol Chem* 284:22898-22904.

88. Nagelhus, T.A., T. Haug, K.K. Singh, K.F. Keshav, F. Skorpen, M. Otterlei, S. Bharati, T. Lindmo, S. Benichou, R. Benarous, and H.E. Krokan. 1997. A sequence in the N-terminal region of human uracil-DNA glycosylase with homology to XPA interacts with the C-terminal part of the 34-kDa subunit of replication protein A. *J Biol Chem* 272:6561-6566.

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

90. Reitmair, A.H., M. Redston, J.C. Cai, T.C. Chuang, M. Bjerknes, H. Cheng, K. Hay, S. Gallinger, B. Bapat, and T.W. Mak. 1996. Spontaneous intestinal carcinomas and skin neoplasms in Msh2-deficient mice. *Cancer Res* 56:3842-3849.

91. Baker, S.M., A.W. Plug, R.A. Prolla, C.E. Bronner, A.C. Harris, X. Yao, D.-M. Christie, C. Monell, N. Arnheim, A. Bradley, T. Ashley, and R.M.

Liskay. 1996. Involvement of mouse *Mlh1* in DNA mismatch repair and meiotic crossing over. *Nature Gen.* 13:336-342.

92. Baar, J., N.M. Pennell, and M.J. Shulman. 1996. Analysis of a hot spot for DNA insertion suggets a mechanism for Ig switch recombination. *J. Immunol.* 157:3430-3435.

93. Edelmann, W., K. Yang, A. Umar, J. Heyer, K. Lau, K. Fan, W. Liedtke, P.E. Cohen, M.F. Kane, J.R. Lipford, N. Yu, G.F. Crouse, J.W. Pollard, T. Kunkel, M. Lipkin, R. Kolodner, and R. Kucherlapati. 1997. Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. *Cell* 91:467-477.

94. Schrader, C.E., W. Edelmann, R. Kucherlapati, and J. Stavnezer. 1999. Reduced isotype switching in splenic B cells from mice deficient in mismatch repair enzymes. *J Exp Med* 190:323-330.

95. Kaminski, D.A., and J. Stavnezer. 2007. Stimuli that enhance IgA class switching increase histone 3 acetylation at S alpha, but poorly stimulate sequential switching from IgG2b. *Eur J Immunol* 37:240-251.

96. Rush, J.S., M. Liu, V.H. Odegard, S. Unniraman, and D.G. Schatz. 2005. Expression of activation-induced cytidine deaminase is regulated by cell division, providing a mechanistic basis for division-linked class switch recombination. *Proc Natl Acad Sci U S A* 102:13242-13247.

97. Bransteitter, R., P. Pham, P. Calabrese, and M.F. Goodman. 2004. Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase. *J Biol Chem* 279:51612-51621.

98. Wilson, T.M., A. Vaisman, S.A. Martomo, P. Sullivan, L. Lan, F. Hanaoka, A. Yasui, R. Woodgate, and P.J. Gearhart. 2005. MSH2-MSH6 stimulates DNA polymerase eta, suggesting a role for A:T mutations in antibody genes. *J Exp Med* 201:637-645.

99. Ehrenstein, M.R., and M.S. Neuberger. 1999. Deficiency in Msh2 affects the efficiency and local sequence specificity of immunoglobulin class-switch recombination: parallels with somatic hypermutation. *Embo J* 18:3484-3490.

100. Shen, H.M., A. Peters, B. Baron, X. Zhu, and U. Storb. 1998. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science* 280:1750-1752.

101. Pasqualucci, L., A. Migliazza, N. Fracchiolla, C. William, A. Neri, L. Baldini, R.S. Chaganti, U. Klein, R. Kuppers, K. Rajewsky, and R. Dalla-Favera. 1998. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc Natl Acad Sci U S A* 95:11816-11821.

102. Liu, M., J.L. Duke, D.J. Richter, C.G. Vinuesa, C.C. Goodnow, S.H. Kleinstein, and D.G. Schatz. 2008. Two levels of protection for the B cell genome during somatic hypermutation. *Nature* 451:841-845.

103. Staszewski, O., R.E. Baker, A.J. Ucher, R. Martier, J. Stavnezer, and J.E.J. Guikema. Activation-induced cytidine deaminase induces reproducible DNA breaks at many non-Ig loci in activated B cells. *Mol Cell*

104. Imai, K., N. Catalan, A. Plebani, L. Marodi, O. Sanal, S. Kumaki, V. Nagendran, P. Wood, C. Glastre, F. Sarrot-Reynauld, O. Hermine, M. Forveille, P. Revy, A. Fischer, and A. Durandy. 2003. Hyper-IgM syndrome type 4 with a B lymphocyte-intrinsic selective deficiency in Ig class-switch recombination. *J Clin Invest* 112:136-142.

105. Min, I.M., C.E. Schrader, J. Vardo, T.M. Luby, N. D'Avirro, J. Stavnezer, and E. Selsing. 2003. The Smu tandem repeat region is critical for Ig isotype switching in the absence of Msh2. *Immunity* 19:515-524.

106. Doi, T., L. Kato, S. Ito, R. Shinkura, M. Wei, H. Nagaoka, J. Wang, and T. Honjo. 2009. The C-terminal region of activation-induced cytidine deaminase is responsible for a recombination function other than DNA cleavage in class switch recombination. *Proc Natl Acad Sci U S A* 106:2758-2763.

107. Yan, C.T., C. Boboila, E.K. Souza, S. Franco, T.R. Hickernell, M. Murphy, S. Gumaste, M. Geyer, A.A. Zarrin, J.P. Manis, K. Rajewsky, and F.W. Alt. 2007. IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* 449:478-482.

108. Stavnezer, J., A. Bjorkman, L. Du, A. Cagigi, and Q. Pan-Hammarstrom. Mapping of Switch Recombination Junctions, a Tool for Studying DNA Repair Pathways during Immunoglobulin Class Switching. *Adv Immunol* 108:45-109.

109. Ehrenstein, M.R., C. Rada, A.M. Jones, C. Milstein, and M.S. Neuberger. 2001. Switch junction sequences in PMS2-deficient mice reveal a microhomology-mediated mechanism of Ig class switch recombination. *Proc Natl Acad Sci U S A* 98:14553-14558.

110. Schrader, C.E., J. Vardo, and J. Stavnezer. 2002. Role for mismatch repair proteins Msh2, Mlh1, and Pms2 in immunoglobulin class switching shown by sequence analysis of recombination junctions. *J Exp Med* 195:367-373.

111. Kracker, S.e.a. In press. Impaired induction of DNA lesions during immunoglobulin class switch recombination in humans influences end-joining repair. . *Proc. Nat. Sci. USA*

112. Simmons, L.A., B.W. Davies, A.D. Grossman, and G.C. Walker. 2008. Beta clamp directs localization of mismatch repair in Bacillus subtilis. *Mol Cell* 29:291-301.

113. Hardeland, U., C. Kunz, F. Focke, M. Szadkowski, and P. Schar. 2007. Cell cycle regulation as a mechanism for functional separation of the apparently redundant uracil DNA glycosylases TDG and UNG2. *Nucleic Acids Res* 35:3859-3867.

114. Hagen, L., B. Kavli, M.M. Sousa, K. Torseth, N.B. Liabakk, O. Sundheim, J. Pena-Diaz, M. Otterlei, O. Horning, O.N. Jensen, H.E. Krokan, and G. Slupphaug. 2008. Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. *EMBO J* 27:51-61.

115. Wang, Y., D. Cortez, P. Yazdi, N. Neff, S.J. Elledge, and J. Qin. 2000. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev* 14:927-939.

116. Dinkelmann, M., E. Spehalski, T. Stoneham, J. Buis, Y. Wu, J.M. Sekiguchi, and D.O. Ferguson. 2009. Multiple functions of MRN in end-joining pathways during isotype class switching. *Nat Struct Mol Biol* 16:808-813.

117. Akbari, M., M. Otterlei, J. Pena-Diaz, P.A. Aas, B. Kavli, N.B. Liabakk, L. Hagen, K. Imai, A. Durandy, G. Slupphaug, and H.E. Krokan. 2004. Repair of U/G and U/A in DNA by UNG2-associated repair complexes takes place predominantly by short-patch repair both in proliferating and growth-arrested cells. *Nucleic Acids Res* 32:5486-5498.

118. Wu, X., and J. Stavnezer. 2007. DNA polymerase beta is able to repair breaks in switch regions and plays an inhibitory role during immunoglobulin class switch recombination. *J Exp Med* 204:1677-1689.

119. Shahi, A., J.H. Lee, Y. Kang, S.H. Lee, J.W. Hyun, I.Y. Chang, J.Y. Jun, and H.J. You. Mismatch-repair protein MSH6 is associated with Ku70 and regulates DNA double-strand break repair. *Nucleic Acids Res*