UNIVERSITY COLLEGE LONDON

REGULATION OF
ACTIVATION INDUCED DEAMINASE

By

SIIM PAUKLIN

A THESIS PRESENTED TOWARDS
THE DEGREE OF DOCTOR OF PHILOSOPHY
2009

Declaration

I, Siim Pauklin, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Siim Pauklin

Clare Hall

June 2009

Abstract

Activation Induced Deaminase (AID) belongs to the protein family of DNA deaminases, which catalyse the deamination of the cytosine residues in single stranded DNA, resulting in the formation of deoxy-uracils. The enzymatic activity of AID is required for the immunoglobulin gene modifications by class switch recombination (CSR), somatic hypermutation (SHM) and gene conversion (iGC). While being essential for antibody diversification, the activity of AID can be harmful for the organism due to its direct mutagenic activities and induction of genomic instability.

This thesis investigates AID regulation both, on the level of gene expression and its interaction partners, and the DNA repair pathways triggered by AID-mediated DNA deamination. Firstly, I have identified estrogen and progesterone as regulators of AID expression. This is achieved via direct binding of estrogen and progesterone receptors to AID promoter. Estrogen leads to an induction of AID expression and increase in AID-mediated downstream pathways – SHM, CSR as well as oncogenic translocations between Ig and c-myc loci. In contrast, progesterone results in a decrease in AID expression and an attenuation of its downstream pathways. Secondly, by generating DT40 cell lines with endogenously tagged AID, we used co-immunoprecipitation and subsequent mass spectrometry for identifying proteins that form a complex with AID in the cytoplasm, nucleoplasm and chromatin. The results of this approach gave us possible insight into the mechanistic process of AID-mediated DNA deamination *in vivo*, suggesting that chromatin bound AID resides in a complex with elongating RNA polymerase II. Thirdly, by expressing AID in meiotic recombination deficient fission yeast and nematode, we have established that a meiotic cell can process a base mismatch, using the base excision repair machinery, to give rise to meiotic recombination. This suggests that meiotic cells can process lesions other than Spo-11 induced DSBs for recombination.

Table of contents

Table of figures

Abbreviations

A	Adenine
Ab	Antibody
ACT	Actinomycin D
AMA	Alpha amanitin
AID	Activation Induced Deaminase
Ag	Antigen
AP	Apurinic/apyrimidinic
APOBEC	Apolipoprotein B-editing catalytic polypeptide-like subunit family
ATM	Ataxia telangiectasia mutated
ATR	Ataxia telangiectasia mutated and Rad3 related
BER	Base excision repair
BSA	Bovine serum albumine
C	Cytosine
C. elegans	Caenorhabditis elegans, nematode, worm
CIP	Calf intestinal phosphatase
C region	Constant region
C_H	Heavy chain constant region
ChIP	Chromatin immunoprecipitation
CSR	Class switch recombination
(c)DNA	(Copy) deoxyribonucleic acid
CT	Circle transcripts
D region	Diversity region
dA	Deoxyadenine
DAPI	4',6-diamidino-2-phenylindole
dC	Deoxycytosine
dG	Deoxyguanine
DMSO	Dimethyl sulphoxide
DNA-PKcs	DNA-dependent protein kinase catalytic subunit
dNTP	Deoxyribonucleotide triphosphate
dRPase	5′-deoxyribophosphodiesterase
DSB	Double-strand break
dT	Deoxythymine
DTT	Dithiothreitol
dU	Deoxyuracil
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
eGFP	Enhanced green fluorescence protein
EMSA	Electro-mobility shift assay
ER	Estrogen receptor
ERE	Estrogen response element
FACS	Florescence activated cell sorter
FCS	Fetal calf serum
FITC	Fluorescein Isothiocyanate
G	Guanine
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GC	Gene conversion

GREB1	Gene regulated by estrogen in breast cancer 1
Gy	Gray
HIGM	Hyper IgM syndrome
HIV	Human immunodeficiency virus
HRP	Horseradish peroxidase
HRT	Hormone replacement therapy
IP	Immunoprecipitation
IPTG	Isopropyl-β-D-thiogalactoside
IFNγ	Interferon gamma
Ig	Immunoglobulin
iGC	Immunoglobulin gene conversion
IR	Ionizing radiation
IL	Interleukin
J	Joules
J_H	Heavy chain joining region
kDa	Kilodalton
LMB	Leptomycin B
LPS	Lipopolysaccharide
MHC	Major histocompatibility complex
MMR	Mismatch repair
MSH2	DNA mismatch repair protein homologue of bacterial MutS
MRE11	Meiotic recombination 11 homologue
(m)RNA	(Messenger) ribonucleic acid
NBS1	Nijmegen breakage syndrome 1
NER	Nucleotide excision repair
NES	Nuclear export signal
NLS	Nuclear localization signal
NFκB	Nuclear factor kappa B
NHEJ	Non-homologous end joining
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase chain reaction
PKA	Protein kinase A
PNK	T4 polynucleotide kinase
PR	Progesterone receptor
PRE	Progesterone response element
PVDF	Polyvinyldifluoride
R	Purine (A or G)
RACK1	Guanine nucleotide-binding protein subunit beta-2-like-1
Rad51	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)
RAG	Recombination-activating gene
RPA	Replication protein A
RNAi	RNA interference
RT	Room temperature
S	Switch region
S. pombe	Schizosaccharomyces pombe, fission yeast
SDS	Sodium dodecyl sulfate
SHM	Somatic hypermutation
SLE	Systemic lupus erythematosus

SSA	Single-strand annealing
ssDNA	Single stranded DNA
T	Thymine
TBE	Tris-Borate-EDTA
TDG	Thymine-DNA glycosylase
TEV	Tobacco Etch Virus
TGFβ	Tranforming growth factor beta
TNFα	Tumor necrosis factor alpha
QRT-PCR	Quantitative real time PCR
U	Uracil
UNG	Uracil-DNA glycosylase (UDG in bacteria)
UV	Ultraviolet light
V	Variable gene segment
VIF	Virion infectivity factor
V_H	Heavy chain variable region
V_L	Light chain variable region
W	A or T
WRC	A/T-A/G-C
wt	Wild-type
Y	Pyrimidine (C or T)
ψV	IgV pseudogene

Acknowledgements

Chapter 1

INTRODUCTION

1.1 Immune cells and immune responses.

The immune system consists of an innate (non-specific) and adaptive (specific) immune system, which forms an integral network composed of cellular and non-cellular components. The predominant cellular components of the nonspecific immune system are macrophages and of the specific immune system are lymphocytes. Lymphocytes can be divided into B-cells and T-cells. In general, B-cells have the function of making immunoglobulins, whereas T-cells have distinctive functions [Reviewed in (Allen et al., 2007)]. Immunity mediated by antibodies is known as humoral immunity, and immunity that is due to cytotoxic T-cells and/or T helper cells is known as cell-mediated immunity (Abbas, 2007). Cytotoxic T-cells kill cells infected with viruses or other pathogens, whereas a subclass of T-cells named T helpers are essential for shaping the immune responses via regulating the activation of B-cells and other cells of the immune system by secreting cytokines [Reviewed in (Felix and Allen, 2007)]. A subclass of T helpers recognizes the antigenic peptide and major histocompatibility (MHC) complex on B-cells and activates them to proliferate and differentiate into antibody producing plasma cells [Reviewed in (Dorner and Radbruch, 2007)]. This interaction between T helpers and B-cells is required for the initiation of antibody diversification and subsequent antibody production by plasma cells. The antibodies secreted by the plasma cell bind to extracellular pathogens and toxins, making them prone to phagocytosis by macrophages, and promote their destruction by other effector cells or by the complement system (Abbas, 2007). In addition to B-cells and T-cells, other cells are involved in innate and adaptive immune responses. Macrophages and dentritic cells engulf and destroy pathogens by phagocytosis. Dendritic cells are phagocytic when they are immature and take up pathogens, however after maturing they initiate adaptive immune responses by acting as antigen-presenting cells to T-cells. Basophils and mast cells act primarily as secretory cells during adaptive immune responses, releasing their histamine and proteoglycan containing granules upon activation by antibody (Abbas, 2007).

1.2 Antibodies.

Antibodies produced by B-cells play a central part in the immune responses of mammals. The antibody molecules consist of four polypeptide chains – two heavy chains and two light chains. The two heavy chains are linked to each other by disulfide bonds and each heavy chain is linked to a light chain by a disulfide bond (Figure 1). In all immunoglobulin molecules, the two heavy chains and the two light chains are identical, giving an antibody molecule two identical antigen-binding sites. Two types of light chain, termed lambda (λ) and kappa (κ), are found in antibodies. Either type of light chain may be found in antibodies of any five major classes (IgM, IgD, IgG, IgA and IgE), but a given immunoglobulin either has lambda or kappa chains (Abbas, 2007). The class, and thus the effector function of an antibody, is defined by the structure of its heavy chain (μ, δ, γ, α, ε), in IgM, IgD, IgG, IgA and IgE, respectively (Abbas, 2007). Each of the four Ig chains has a variable (V) region at its amino terminus, which contributes to the antigen-binding site, and a constant (C) region, which determines the isotype. Ig isotypes have different functions in the immune system. IgM plays an important role in the activation of the complement system. IgG isotypes are primarily involved in opsonisation, neutralization and activation of the complement system, as well as the sensitization for killing by natural killer cells. IgA is involved in neutralization, and IgE is important for sensitizing mast cells (Abbas, 2007).

The diversity of the immunoglobulin repertoire in humans is generated by several processes (Abbas, 2007). Firstly, the multiple inherited gene segments V, D and J that form the Ig V region, can be used in different combinations. Secondly, variable addition and substraction of nucleotides at the junctions between gene segments increase diversity. Thirdly, different possible combinations of a heavy and light chain are used in the complete immunoglobulins. Fourthly, V regions are mutated in mature B-cells through the process of somatic hypermutation. Somatic hypermutation is triggered by enzymatic deamination of cytosines in the immunoglobulin locus. The enzyme Activation Induced Deaminase that catalyzes this process, and is the main subject of this thesis, will be described in detail below.
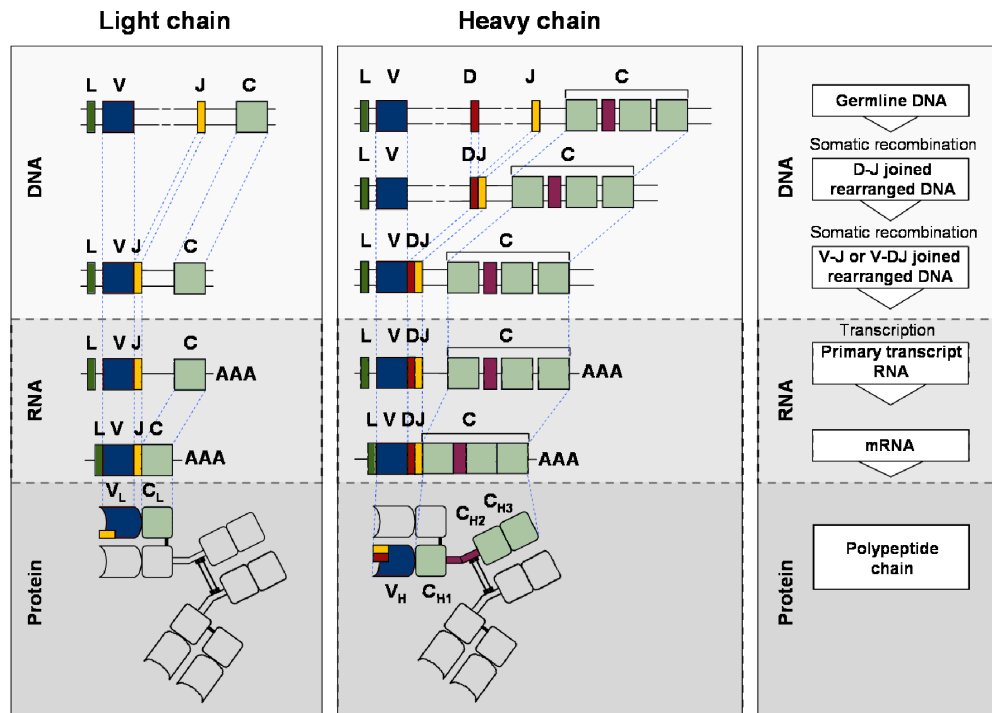
Figure 1. Antibody heavy and light chains are constructed from gene segments. Light-chain V-region genes are constructed from two segments (left pannel). A variable (V) and a joining (J) gene segment in the genomic DNA are joined to form a complete light-chain V-region. A leader peptide (L) directs the protein into the cell's secretory pathways and is then cleaved. The light-chain C region is joined to the V-region by splicing of the light-chain RNA to remove the L-to-V and J-to-C introns. Heavy-chain V regions are constructed from three gene segments (centre panel). First, the diversity (D) and J gene segments join, followed by the joining of V segment with DJ the combined DJ sequence, forming a complete VH. The C-region, which consists of several segments, is spliced to the V-domain sequence together with the leader sequence. The leader sequence is removed after translation and the disulfide bonds (black lines) that link the polypeptide chains are formed. The hinge region is shown in purple. Adapted from (Janeway, 2001).

1.3 Activation Induced Deaminase.

Activation Induced Deaminase is a 24 kDa protein that is a member of the AID/APOBEC family of proteins (Figure 2) [Reviewed in (Petersen-Mahrt, 2005)]. AID is produced by mature vertebrate B-lymphocytes that have been activated to undergo somatic hypermutation and class switch recombination. It is also found in immature B-cells of birds and some mammalian species undergoing gene conversion of Ig genes [Reviewed in (Di Noia and Neuberger, 2007)]. AID enzymatic activity is essential for these processes, indicated by the hyper IgM syndrome in human patients lacking AID and in AID knock-out mice

(Muramatsu et al., 2000; Revy et al., 2000). It has been shown *in vitro* (Bransteitter et al., 2003; Chaudhuri et al., 2003; Dickerson et al., 2003) and in *E. coli* (Petersen-Mahrt et al., 2002), that AID initiates these events by deaminating cytosines in DNA into uracils (U). AID preferentially targets WRC motifs (A/T-A/G-C) in ssDNA *in vitro* (Yu et al., 2004a), in yeast (Mayorov et al., 2005) and reveals hotspots for nucleotide mutations during somatic hypermutation and in
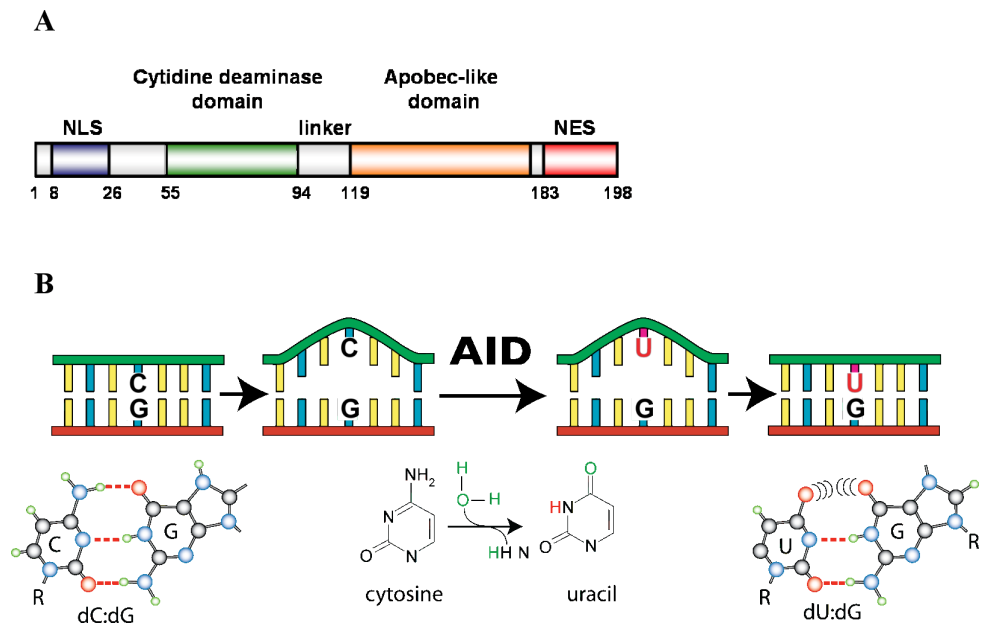
Figure 2. Activation induced deaminase (AID). (A) The structural domains of AID. NLS – nuclear localization signal, NES – nuclear export signal. Numbers indicate the corresponding amino acids. (B) The catalytic activity of AID in ssDNA produces a uracil from a cytosine by deamination. This results in a U:G mismatch in DNA. Modified from (Durandy et al., 2006; Neuberger et al., 2003).

immunoglobulin switch regions (Larijani et al., 2005; Shapiro and Wysocki, 2002). Not much is known about the mechanisms of AID's targeting to the Ig locus and the molecular reasons for its selective activity on these loci. There are a number of genes that are mutated by AID such as *BCL6, CD95/FAS, c-myc, RHO/TTF, PAX-5, PIM1, Igα, Igβ, B29* and *MB1* (Gordon et al., 2003; Kuppers and Dalla-Favera, 2001; Liu et al., 2008; Muschen et al., 2000; Pasqualucci et al., 1998; Pasqualucci et al., 2001; Shen et al., 1998). The difference in AID-mediated mutations in various loci appears to be in part due to regulated targeting of AID, as overexpression of AID causes mutation of several genes that in normal B-cells are not subject to somatic hypermutation (Martin et al., 2002;

Okazaki et al., 2003; Yoshikawa et al., 2002). In addition to specific targeting of AID to certain loci, a balance between high fidelity DNA repair and error-prone repair seems to determine whether a gene will accumulate AID-induced mutations or not (Liu et al., 2008).

Cis elements, including enhancers, have been suggested to affect AID targeting to certain loci. The presence of two additional E2A protein-binding sequences CAGGTG has been shown to increase V region mutation frequencies (Michael et al., 2003). The oncogene c-myc, that also contains this relatively frequently occurring cis element, is expressed and mutated in B-cell lymphomas (Klein et al., 2003; Pasqualucci et al., 2001). Somatic hypermutation of Igκ is also affected by intronic or 3' enhancers, but this is probably not the only determinant for the efficiency of mutations, because it occurs in mice that lack the intronic enhancer (Betz et al., 1994; Perlot et al., 2005; van der Stoep et al., 1998).

Most of AID protein is localized in the cytoplasm of B-cells (Ito et al., 2004; McBride et al., 2006). AID has an N-terminal nuclear import signal as well as a C-terminal nuclear export signal and accumulates in the nucleus after overexpression in cells where nuclear export is inhibited (McBride et al., 2006). The half-life of AID protein is shorter in the nucleus compared to the cytoplasm, due to its polyubiquitinylation and faster degradation in the nucleus (Aoufouchi et al., 2008). In addition, the localization and activity of AID are potentially regulated by post-translational modifications. AID has been shown to associate with PKA in a cAMP dependent manner in the cytoplasm of Ramos cells (Pasqualucci et al., 2006), and its phosporylation has been correlated with AID localization to the nucleus. Purified AID from activated mouse B-cells is phosphorylated at serine 38 (S38), tyrosine 184 (Y184) (Basu et al., 2005), and threonine 140 (T140) (McBride et al., 2008). Threonine 140 and serine 38 can also be phosphorylated *in vitro* by various PKC isoforms, and the mutation of these residues in AID protein leads to a decrease in somatic hypermutation and also in Ig class switching (McBride et al., 2008). AID purified from HEK293 non-lymphoid cells is not as highly phosphorylated and lacks substantial interaction with RPA (Chaudhuri et al., 2004), a protein that has been identified to bind AID. However, phosphorylation of AID *in vitro* renders the protein capable of interacting with RPA (Basu et al., 2005). In addition to RPA, AID

may interact with Mdm2 (MacDuff et al., 2006), the transcription machinery (Nambu et al., 2003), and CTNNBL1 [a nuclear protein that interacts with the spliceosome (Conticello et al., 2008)]. AID mutants that are unable to bind CTNNBL1 yield severely diminished hypermutation and class switching and the absence of CTNNBL1 in DT40 results in abrogated Ig V diversification (Conticello et al., 2008).

The regulation of AID expression in B-cells is mediated by both stimulatory and inhibitory factors. AID expression is induced by CD40, bacterial lipopolysaccharides and interleukin-4 (IL-4) (produced by CD4+ T-cells) (Muramatsu et al., 1999). These signals are mediated by various transcription factors such as Pax5, E47, Stat6, NFκB, Irf8 and Hox4C, all of which function in a cooperative manner with negative regulators of AID such as id3 or Blimp1 – classical B-cell specific transcription factors (Dedeoglu et al., 2004; Gonda et al., 2003; Lee et al., 2006; Sayegh et al., 2003; Shapiro-Shelef et al., 2005; Xu et al., 2007). Furthermore, AID is negatively regulated by microRNA-155 as well as microRNA-181b in a posttranscriptional manner (de Yebenes et al., 2008; Dorsett et al., 2008; Teng et al., 2008).

1.4 Somatic hypermutation.

Somatic hypermutation introduces point mutations in rearranged V region genes, generating antibodies with higher affinity to antigens (Figure 3) (Weigert et al., 1970; Neuberger et al., 2003). AID induced mutant immunoglobulin molecules that bind antigen better than the original B-cell receptors are preferentially selected to mature into antibody-secreting cells. This process is called antibody affinity maturation. The mutations introduced by SHM are predominately point mutations, although insertions and deletions are occasionally observed (Goossens et al., 1998).

The U:G mismatches due to AID have several possible fates in SHM (Neuberger et al., 2003; Odegard and Schatz, 2006). If a lesion is not repaired before the start of DNA replication, DNA polymerases will insert an A nucleotide opposite the U nucleotide creating C to T (G to A) transition mutations. If the U nucleotide is removed by uracil-DNA glycosylase (UNG), an abasic site is

created, replication of which should give rise to both transition (purine to another purine) and transversion (purine to pyrimidine) mutations. In addition to activating UNG-dependent base-excision repair (BER), a U:G mismatch can also recruit the mismatch repair (MMR) machinery (Wilson et al., 2005), which allows for mutations at A:T near the initiating U:G lesion through an error-prone patch repair process by DNA polymerase eta (Polη) (Delbos et al., 2007; Delbos et al., 2005).
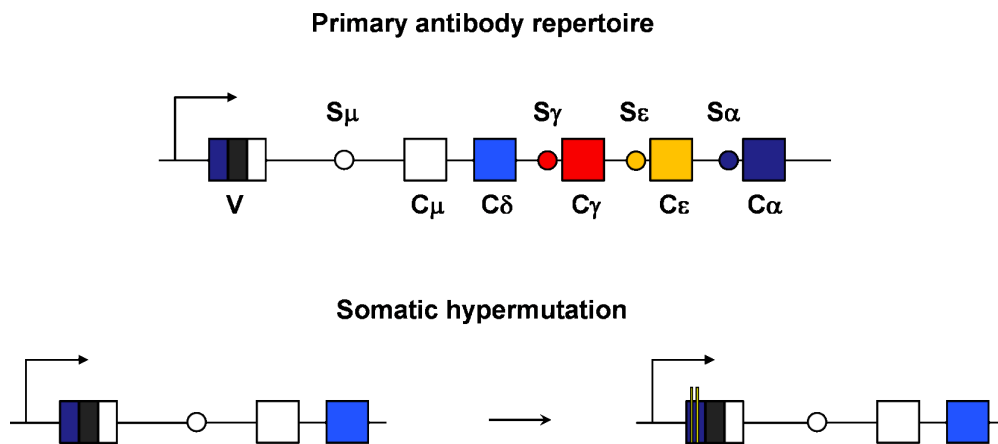
Figure 3. Schematic drawing of somatic hypermutation during affinity maturation of antibodies. A rearranged Ig gene as would be present in mature naive B-cells is depicted at the top. The V segment encodes the Ig variable region whereas the C regions encode constant regions. *S* denotes switch regions. The arrow preceding the V segment represents a promoter. SHM introduces non-templated point mutations in V genes. Mutations are shown as yellow vertical lines in the V segment.

Transcription of the immunoglobulin locus is essential for SHM, as loss of transcription abolishes the frequency of mutations, and the mutation rate of a reporter immunoglobulin gene is proportional to the rate of transcription through that locus (Bachl et al., 2001; Fukita et al., 1998). Mutations are confined to a 1–2 kilobase (kb) region in rearranged immunoglobulin genes, beginning about 150 base pairs (bp) downstream of the transcription start site and the frequency of mutations decreases exponentially with increasing distance from the transcription start site {Lebecque, 1990 #678}(Rada and Milstein, 2001).

U nucleotides are generated during normal cell growth due to spontaneous deamination of cytosines and are misincorporated into DNA during replication

(Krokan et al., 2002). These U nucleotides are efficiently and accurately repaired, predominately by BER pathways that depend on the activity of uracil-DNA glycosylases. During SHM, however, the uracil lesion is repaired in an error-prone manner, resulting in an approximate mutation rate of $2x10^{-3}$ mutations per base pair per cell division, which is 10^6-fold higher than the spontaneous mutation rate in somatic cells (Odegard and Schatz, 2006). One mechanism by which error-free repair might be disrupted, involves MRE11 (meiotic recombination 11 homologue)–RAD50 complex, which is able to cleave DNA at abasic sites, generating a 3' end that cannot be extended by DNA polymerases (Larson et al., 2005a). Another study, in chicken DT40 cells, indicated the importance of PCNA ubiquitinylation in somatic hypermutation, presumably by recruiting error-prone DNA polymerases (Arakawa et al., 2006).

1.5 Class switch recombination.

Class switch recombination detaches an expressed heavy chain variable (VDJ) region from one constant (C) region and joins it to another downstream C region, deleting the DNA in between (Figure 4) [Reviewed in (Chaudhuri and Alt, 2004)]. Switch recombination rejoins the chromosome and generates switch circles carrying the excised fragment. The ectopic expression of AID is sufficient to induce switch recombination (Okazaki et al., 2002) and somatic hypermutation (Yoshikawa et al., 2002) in engineered minigenes in mammalian fibroblasts,
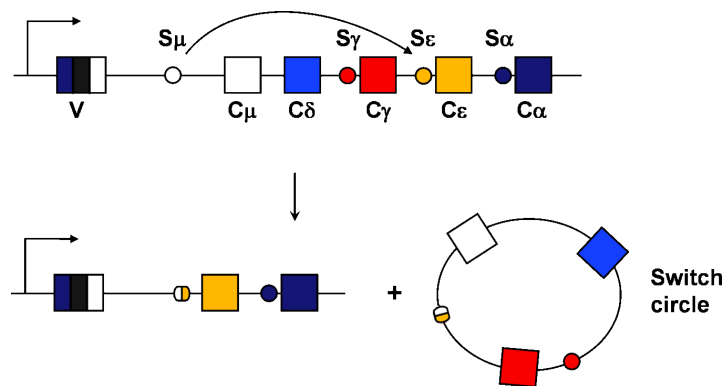
Figure 4. A schematic drawing of class switch recombination. CSR occurs between a Sμ region and a S region located upstream of another C region, such as Sε, which will result in the switching from IgM to IgE. The intervening DNA sequence is predominantly released as a circular episome called a switch circle.

indicating that downstream steps in the AID-initiated pathways can be carried out by non-B-cell specific factors. Recombination is region-specific, not sequence-specific, and produces junctions at heterogeneous sites and sequences within the switch (S) regions, which are repetitive and degenerate guanine-rich regions 2–10 kb in length (Dunnick et al., 1993). Extracellular signals delivered by T-cells and cytokines, which activate S region transcription via an enhancer and promoter in the intron upstream of each S region, regulate the production of specific classes of Ig. Thus, the critical signal for recombination is transcription of the S regions targeted for recombination (Shinkura et al., 2003). Recently switch regions (not constant regions) have been reported to be transcribed in anti-sense direction (similarly to V region), but the exact role of this unusual transcription event remains unclear (Perlot et al., 2008).

CSR depends upon a 10 amino acid region at the very C terminus of AID, a region unnecessary for cytosine deamination, gene conversion or somatic hypermutation (Barreto et al., 2003; Shinkura et al., 2004; Ta et al., 2003). Targeted disruption of *ung*, *Msh2*, *Msh6*, and *Exo1* in mice decrease switch recombination, indicating the involvement of these proteins in the process (Rada et al., 2004; Rada et al., 2002). CSR also depends upon factors involved in nonhomologous end-joining, including Ku and DNA-PKcs (Casellas et al., 1998). Furthermore, CSR requires four factors involved in the DNA damage response: the cell cycle regulator ATM (Lumsden et al., 2004; Reina-San-Martin et al., 2004); the phosphorylated variant histone γ-H2AX (Reina-San-Martin et al., 2003); NBS1, the regulatory component of the MRE11/RAD50/NBS1 (MRN) complex (Manis et al., 2004); and the p53-binding protein 1, 53BP1 (Reina-San-Martin et al., 2004; Reina-San-Martin et al., 2003; Ward et al., 2004). These factors act as early sensors of AID-mediated DNA lesions and coordinate downstream DNA repair pathways leading to class switch recombination.

1.6 Gene conversion.

Gene conversion (GC) mediates the transfer of genetic information from intact homologous sequences to the region that contains a DSB – this can occur between sister chromatids, homologous chromosomes or homologous sequences

on either the same chromatid or different chromosomes [Reviewed in (Chen et al., 2007)]. Immunoglobulin GC (iGC) is the process through which immunoglobulins are diversified in chicken, rabbit, cattle and pigs (Figure 5) (Arakawa and Buerstedde, 2004; Reynaud et al., 1985; Reynaud et al., 1987). This process introduces sequences from adjacent V pseudogenes into the expressed V gene, creating diversity in the receptors (Reynaud et al., 1985; Reynaud et al., 1987). In chicken, all B-cells express the same surface immunoglobulin (sIg) initially; there are several D segments but only one active V and J gene segment for the chicken heavy-chain gene and one active V and J gene segment for each of the light-chain genes. Gene rearrangement in chicken can thus produce only a single receptor specificity (Reynaud et al., 1985; Reynaud et al., 1987).

The chicken DT40 cell line has been a useful tool for studying iGC and the factors involved in this process. iGC is initiated by deamination of cytosines by AID, followed by base removal by UNG, which creates an abasic site that is prone to cleavage (Arakawa et al., 2002; Arakawa et al., 2004; Harris et al., 2002b). The resulting DNA 3' end undergoes homologous repair, promoted by RAD51 paralogs XRCC2, XRCC3 or RAD51B and BRCA2, copying new sequence information from an upstream donor pseudogene (Krogh and Symington, 2004; Sale et al., 2001). It has also been suggested that single-strand lesions lead to gene conversion via Nbs1-mediated homologous recombination (Nakahara et al., 2009).
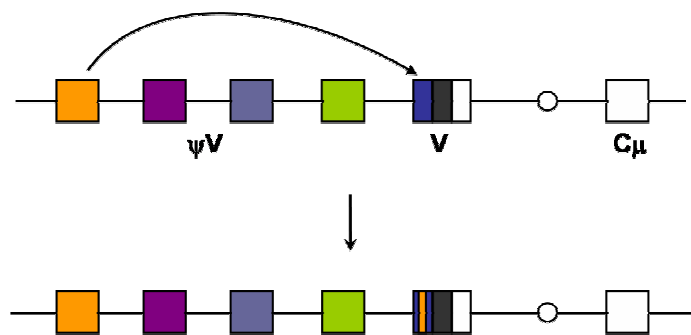
Figure 5. A schematic drawing of gene conversion. Gene conversion introduces mutations in V genes by shuffling sequences from the upstream pseudogenes (ψV, coloured boxes) with similar sequences in the functional V region. The ψV gene acting as a template is shown in orange.

In summary, AID-mediated deamination of cytosines to uracils is the first step that can trigger a number of downstream DNA repair pathways, which can result in the simple repair of the lesion, mutations in or near the lesion or recombinational events such as gene conversion or class switch recombination (Figure 6).
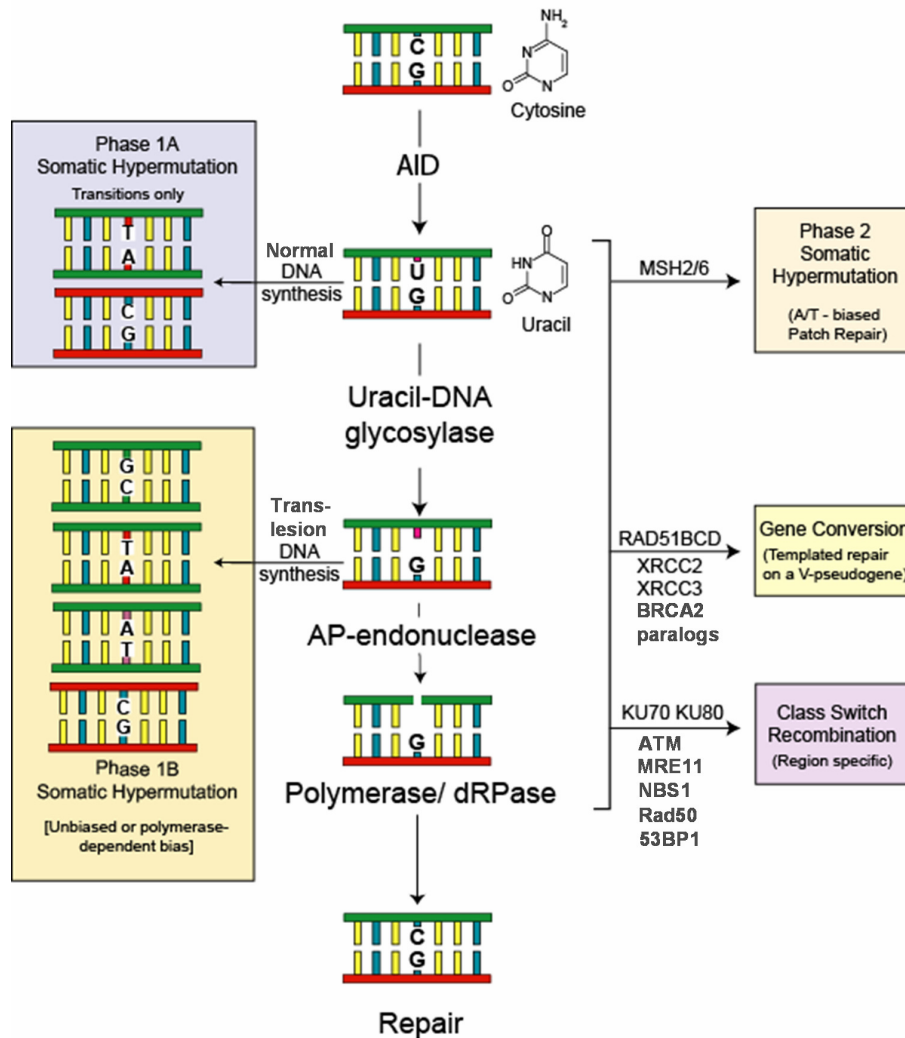
Figure 6. Differential processing of AID-induced lesions leads to SHM, iGC or CSR. AID deaminates cytosines in the Ig locus leading to U:G mismatches that can be processed by Uracil-DNA glycosylase (UNG) or mismatch repair proteins MSH2/6. These DNA repair pathways cooperate with distinct sets of DNA repair factors that are involved in processing AID-induced lesions into mutations for SHM, V-pseudogene templated gene conversion or DSB repair for CSR. SHM process can be divided into several phases (Phase 1A, 1B or Phase 2), depending on the mechanism of the mutation generation process. Modified from (Neuberger et al., 2003).

1.7 APOBEC family members.

In addition to AID, the AID/APOBEC family (comprising of APOBEC1, APOBEC2, APOBEC3(s) and APOBEC4) contains members that can deaminate cytosine in DNA and/or RNA and exhibit diverse physiological functions [Reviewed in (Chiu and Greene, 2008; Harris and Liddament, 2004)]. APOBEC1 is an RNA cytidine deaminase, which edits apolipoprotein B RNA, deaminating C^{6666} to U, thereby creating a premature stop codon and potentiating the tissue-specific production of a truncated apolipoprotein B involved in lipid metabolism (Navaratnam et al., 1993; Teng et al., 1993). On the other hand, APOBEC2 protein shows a muscle specific expression, but its function, similarly to APOBEC4, remains unknown (Liao et al., 1999; Rogozin et al., 2005). However, Apobec2 and Apobec3 double knock-out mice reveal no obvious phenotypes, suggesting that these genes are not essential for the viability or fertility of the organism (Mikl et al., 2005), although Apobec3 has been shown to inhibit the movement of retrotransposable elements (Chiu and Greene, 2008).

APOBEC3 subgroup of the family has undergone a significant expansion in higher mammals, especially in primates (Figure 7) (Conticello et al., 2005). Mouse has only one Apobec3 gene, whereas the human genome encodes at least eight APOBEC3 members, named APOBEC3A-H [Reviewed in (Holmes et al., 2007b)]. Several APOBEC3 proteins have been shown to function as cellular factors for restricting retroviral infection. APOBEC3G protein is included during the packaging of the HIV virion and carried with newly produced viral particles to a new target cell (Sheehy et al., 2002). Upon infection of the new cell, APOBEC3 protein triggers the deamination of cytosine residues to uracils in the first cDNA strand of the replicating retroviral genome (Harris et al., 2003). One function of the uracils is as a template for the incorporation of second-strand adenines, thereby fixing the lesions as G to A hypermutations. This process leads to a detrimental level of mutations in the viral genome, decreasing the viability of the virus. In addition, APOBEC proteins can inhibit retroviral infections by a deamination independent mechanism that likely involves the direct inhibition of the viral reverse transcriptase (Bishop et al., 2006; Chiu and Greene, 2008; Newman et al., 2005). Besides APOBEC3G, anti-retroviral (e.g HIV) functions have also been attributed to several other APOBEC3 proteins, such as human

APOBEC3B and APOBEC3F and non-human APOBEC proteins like rat Apobec1 and mouse Apobec3 (Holmes et al., 2007a). APOBEC3G is in a large complex with ribonucleoproteins, found in cytoplasmic structures called P-bodies and stress granules (Chiu et al., 2006; Gallois-Montbrun et al., 2007; Kozak et al., 2006). The importance of APOBEC3 mediated retroviral restriction is substantiated by the presence of *Vif* (virion infectivity factor) gene in the genomes of human and rodent immunodeficiency viruses. VIF protein leads to the degradation of APOBEC3 protein, thereby decreasing the pool of this protein (Sheehy et al., 2002). This disables the cellular defence mechanisms against viral infections (Yu et al., 2003).
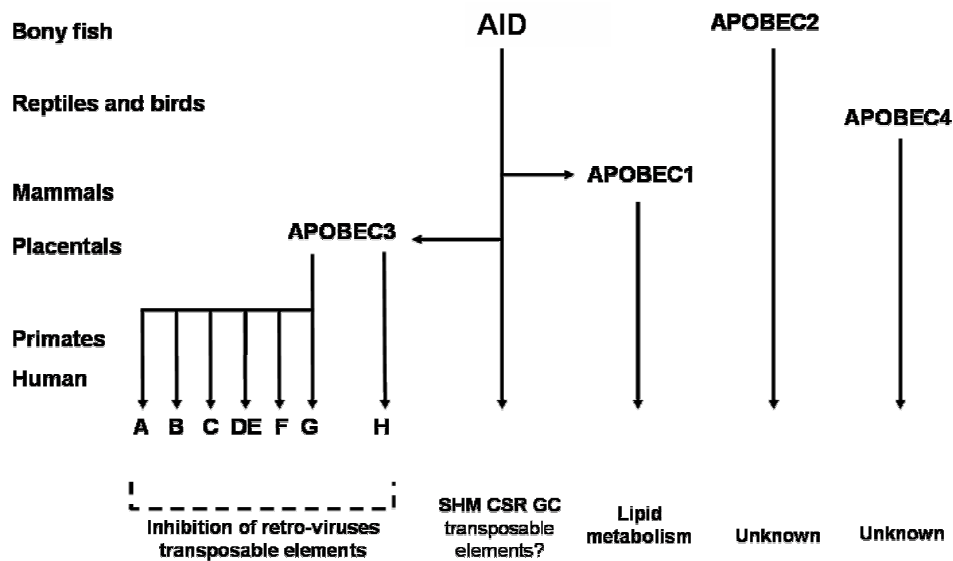
Figure 7. Phylogenetic relationship of the APOBEC gene family. AID and APOBEC2 are the ancestral members of the family, found in primates, birds as well as bony fish. APOBEC1 and the APOBEC3 subfamily (APOBEC3A to 3H) have later evolved from AID via gene duplications in mammals. This has culminated in primates and human, which have at least 12 APOBEC family members. APOBEC3A-H inhibit retroviral infection and the spreading of transposable elements. AID is required for somatic hypermutation, class switch recombination and gene conversion. Additionally, it may also limit the propagation of transposable elements. APOBEC1 is involved in lipid metabolism. The origin and function of APOBEC2 and APOBEC4 is unknown. Adapted from (Conticello et al., 2007).

In addition to fighting off exogenous retroviral infections, APOBEC3 proteins also block the movement of endogenous retrotransposable elements such as long interspersed nuclear element 1, short interspersed nuclear elements and

Alu sequences (Muckenfuss et al., 2006; Stenglein and Harris, 2006). Retroelements have a common step in their life cycle, the self-replication via an RNA transcript that is copied into single-stranded DNA by reverse transcription, which become the target for APOBEC3 proteins (Esnault et al., 2005).

Evolutionarily, AID/APOBEC family is found in vertebrates with homologs of AID and of APOBEC2 identifiable in bony fish, birds, amphibians, and mammals (Figure 7) (Conticello et al., 2005). Thus, AID and APOBEC2 are likely to be the ancestral members of the AID/APOBEC family and both APOBEC1 and APOBEC3 are mammal-specific derivatives of AID. Whether AID stems from APOBEC2 or *vice versa*, is still unknown.

1.8 DNA deamination and cancer.

The spontaneous deamination of cytosine (as well as of 5-methyl cytosine) in the genome can contribute to both somatic and germline mutations [Reviewed in (Lindahl, 1993)]. A similar danger of mutations arises from the enzymatic activities of DNA deaminases. Indeed, mistargeted SHM by AID has been shown to contribute to lymphomagenesis, by mutating regulatory and coding sequences in proto-oncogenes and tumor suppressor genes, and by promoting chromosomal translocations (Pasqualucci et al., 2001; Ramiro et al., 2004). Many B-cell malignancies harbor chromosomal translocations in which proto-oncogenes are rearranged into the Ig loci. In a high proportion of Burkitt's lymphomas, the breakpoints of translocations in the Ig heavy chain locus occur within one of the switch regions, which strongly suggests, that they are the products of aberrant switch recombination (Rabbitts et al., 1984). For instance, the c-*myc* proto-oncogene undergoes both aberrant hypermutation and translocation in B-cell tumors (Pasqualucci et al., 2001). Moreover, AID is required for c-*myc* translocation leading to tumorigenesis in a murine model for Burkitt's lymphoma (Ramiro et al., 2004; Robbiani et al., 2008). Even though electron microscopic imaging has revealed that G-loops form within the transcribed c-*myc* gene, with structures similar to those that form in transcribed S regions (Duquette et al., 2005), and that these G-loops can act as targets for AID (Duquette et al., 2005), the mechanism for AID-induced translocation *in vivo* is unknown. In addition to

c-*myc*, a small subset of other non-Ig genes undergo aberrant hypermutation or translocation in B-cells, including the *BCL6*, *CD95/FAS*, *RHO/TTF*, *PAX-5,* and *PIM1* proto-oncogenes (Kuppers and Dalla-Favera, 2001; Muschen et al., 2000; Pasqualucci et al., 2001); and two non-oncogenic B-cell receptor genes *B29* and *MB1* (Gordon et al., 2003). The mutation frequency of non-immunoglobulin genes is ~50–1000 fold lower than that of hypermutated variable regions, with the reason for this difference remaining elusive. The acquired mutations resemble those introduced by SHM in rearranged variable regions, as they are predominately transition mutations and are biased to WRC hotspots. Furthermore, the mutation distribution in non-immunoglobulin genes is similar to that in hypermutated variable regions (Gordon et al., 2003; Kuppers and Dalla-Favera, 2001; Muschen et al., 2000; Pasqualucci et al., 2001).

Artificial retroviral substrates lacking any immunoglobulin element mutate when integrated at many positions in the genome of hypermutating B-cell lines (Wang et al., 2004) or in Ramos (Parsa et al., 2007). Some insertion sites revealed a high mutation rate compared to others that had a low mutation rate. This would suggest, that AID-mediated SHM can occur in many genomic locations, and that targeting of SHM to immunoglobulin loci might be less specific than previously thought. This is supported by the finding that a balance between high fidelity DNA repair and error-prone repair determines whether a gene will accumulate AID-induced mutations or not (Liu et al., 2008).

Interestingly, in addition to B-cell lymphoma (Pasqualucci et al., 2008) and multiple myeloma (Chesi et al., 2008), unregulated expression of AID can induce various non-B-cell cancers, such as lung and T-cell (Okazaki et al., 2003), liver (Endo et al., 2007) and gastric cancer (Matsumoto et al., 2007). These gastric cancers have been reported to be triggered by *Helicobacter pylori* infections and are mediated by NFκB dependent upregulation of AID (Matsumoto et al., 2007). It is possible that tissues or cell types other than B-cells, might be more susceptible to AID induced oncogenesis compared to B-cells, in which AID expression is naturally high. If AID is transiently expressed in these tissues, it may result in translocations and mutations that could give rise to uncontrolled proliferation and survival of the cells. Thus, AID might be involved in various different malignancies besides lymphocytes, and the mechanisms are just starting to be elucidated.

Chapter 2

AIMS

2.1 AIMS.

AID is essential for normal humoral immune responses via initiating antibody diversification. By deaminating cytosines in ssDNA at the immunoglobulin loci, AID triggers somatic hypermutation, class switch recombination or gene conversion. On the other hand, if mis-regulated, AID can contribute to some B-cell mediated autoimmune diseases or promote cancer.

In order to investigate how AID is regulated in the cell, both on the level of gene expression as well as its potential interaction partners, I used three approaches:

1. Since humoral immune responses are a target for hormonal regulation, and AID expression has been found in some hormone responsive tissues, I examined the possibility that AID expression could be affected by hormones estrogen and progesterone.

2. In order to glean insight to the network of proteins that form a complex with AID, and therefore modulate its function in B-cells, I aimed to study AID localization and to identify AID-interacting proteins in an unbiased manner using endogenously tagged AID.

3. To set up a system for investigating DNA repair pathways that are utilized by the cell to fix dU lesions, I aimed to study the consequences of AID activity in two model organisms - the fission yeast and the nematode. To analyse DNA repair events initiated from AID-induced dU lesions, I aimed to use AID expressing fission yeast and nematode compromised for meiotic recombination.

Chapter 3

MATERIALS AND METHODS

3.1 Materials.

- All chemicals in materials section are from Sigma except where indicated.
- All media or buffers marked with "*" were prepared by CRUK media kitchen.

3.1.1 ChIP dilution buffer - 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl (pH 8.1), 167 mM NaCl.

3.1.2 Coupling buffer - 0.1M $NaHCO_3$ (pH 8.3), 0.5 M NaCl.

3.1.3 De-staining solution - 10% acetic acid, 40% ethanol.

3.1.4 Dilution and Equilibration Buffer - 20 mM HEPES (pH 7.7), with 1.5 mM $MgCl_2$, 0.2 mM EDTA, 10 mM KCl and 25% Glycerol.

3.1.5 DT40 Lysis Buffer - 50 mM TrisHCl (pH 7.5), 100 mM NaCl, 50 mM KCl, 2 mM $MgCl_2$, 1mM EDTA, 10% Glycerol, 1 tablet of Complete Protease Inhibitors (Roche, UK) per 50 ml, 1 Phosphatase Inhibitor Cocktail Tablet (Roche, UK) per 10 ml.

3.1.6 DT40 Lysis Buffer+T - DT40 lysis buffer + 0.3% Triton X-100.

3.1.7 DT40 Lysis Buffer+TB - DT40 lysis buffer+T + 50 U/ml Benzonase (VWR International, UK).

3.1.8 DT40 Lysis Buffer+TD - DT40 lysis buffer+T + 25 U/ml DNAse I (Roche, UK).

3.1.9 DT40 Lysis Buffer+TDR - DT40 lysis buffer+TD + 0.5 mg/ml RNase A (Sigma, UK).

3.1.10 Elution buffer - DT40 Lysis Buffer + 500 µg/ml 3xFLAG peptide.

3.1.11 EMM: Edinburgh minimal medium* - 3 g/l potassium hydrogen phthalate (final concentration 14.7 mM), 2.2 g/l Na2HPO$_4$ (15.5 mM), 5 g/l NH$_4$Cl (93.5 mM), 20 g/l glucose (2%); 20 ml/l salt solution [52.5 g/l MgCl$_2$*6H$_2$O (0.26 M), 0.735 g/l CaCl$_2$*2H$_2$O (4.99 mM), 50 g/l KCl (0.67 M), 2 g/l Na$_2$SO$_4$ (14.l mM)]; 1 ml/l vitamins [1 g/l pantothenic acid (4.20 mM), 10 g/l nicotinic acid (8l.2 mM), 10 g/l inositol (55.5 mM), 10 mg/l biotin (40.8 µM)]; 0.1 ml/l minerals [5 g/l boric acid (80.9 mM), 4 g/l MnSO$_4$ (23.7 mM), 4 g/l ZnSO$_4$.7H$_2$O (13.9 mM), 2 g/l FeCl$_2$*6H$_2$O (7.40 mM), 0.4 g/l molybdic acid (2.47 mM), 1 g/l KI (6.02 mM), 0.4 g/l CuSO$_4$*5H$_2$O (1.60 mM), 10 g/l citric acid (47.6 mM)]. The media was supplemented with 225 mg/l amino acids (adenine, leucine, histidine, lysine, uracil) according to the experimental requirements. Solid media was made by adding 2% Difco Bacto Agar.

3.1.12 EMSA binding buffer - 20 mM HEPES (pH 7.8), 50 mM KCl, 10% glycerol, 0.25 mM DTT, 0.1 mM EDTA, 0.55 µg of poly(dI-dC) (GE Healthcare, UK) and 0.25 pmol of labeled oligonucleotides.

3.1.13 Extraction buffer - 20 mM HEPES (pH 7.7), 1.5 mM MgCl$_2$, 0.42 M NaCl, 0.2 mM NaCl, 0.2 mM EDTA and 25% glycerol supplemented with Complete Protease Inhibitors and 1 mM DTT.

3.1.14 High salt immune complex wash buffer - 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl (pH 8.1), 500 mM NaCl.

3.1.15 Isotonic lysis buffer - 10 mM TrisHCl (pH 7.5), 2 mM MgCl$_2$, 3 mM CaCl$_2$, and 0.32 M sucrose supplemented with 1 Protease Inhibitor Cocktail Tablet per 50 ml.

3.1.16 5xKCM - 500 mM KCl, 150 mM CaCl$_2$, 250 mM MgCl$_2$.

3.1.17 LB-agar plates* - 1% peptone, 0.5% yeast extract, 0.5% NaCl, 1.2% bacto agar, 50 mg/L selection marker (pH 7.0).

3.1.18 LiCl immune complex wash buffer - 0.25M LiCl, 1% IGEPAL CA630, 1% deoxycholic acid, 1 mM EDTA, 10 mM Tris-HCl (pH 8.1), TE buffer [10 mM Tris-HCl (pH 8.0), 1 mM EDTA].

3.1.19 Low salt immune complex wash buffer - 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl (pH 8.1), 150 mM NaCl.

3.1.20 Lysis buffer for single worm PCR - 50 mM KCl, 10 mM Tris pH 8.3, 2.5 mM MgCl2, 0.45% IGEPAL CA 630, 0.45% Tween-20, 0.01% Gelatin with proteinase K (1 mg/ml).

3.1.21 6 x loading dye - 1.3 mM bromophenol blue, 1.7 mM xylene cyanol FF, 60% glycerol, 5 mM EDTA.

3.1.22 Myo B medium* - 500 g Bactotryptone was mixed with 217 g NaCl, 60 g TrisHCl, 26 g Tris (base) and 6.9 g Cholesterol. 7.2 g from the mixture was used per 1 liter of medium, supplemented with 20 g agar.

3.1.23 NGM Classic - NGM Classic Agarose plates seeded with RNAi expressing bacteria were used for RNA interference (RNAi) experiments for nematode. 12 g NaCl, 64 g Agar (Difco Bacto-Agar) and 10 g Bacto-peptone were dissolved in 4 L of water, followed by autoclaving. This was allowed to cool down to 60 ºC and then, 4 ml cholesterol (5 mg/ml in 95% ethanol), 4 ml 1M $MgSO_4$, 4 ml 1M $CaCl_2$, 100 ml HPO_4 (pH 6.0), Isopropyl β-D-1-thiogalactopyranoside (IPTG; final concentration 6 mM) and plasmid antibiotics (ampicillin 25 µg/ml) were added to the solution and mixed. Approximately 10 ml NGM Classic media was poured into each Petri dish and stored at 4 ºC 1-3 days prior to seeding with RNAi expressing bacteria.

3.1.24 PBS buffer* - 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 140 mM NaCl (pH 7.4).

3.1.25 5 x SDS sample buffer - 2% SDS, 50% glycerol, 0.2% bromophenol blue, 40 mM DTT, 250 mM Tris (pH 6.8).

3.1.26 SDS lysis buffer for ChIP - SDS 1%, 10 mM EDTA, 50 mM Tris-HCl (pH 8.1).

3.1.27 Staining solution - Coomassie Brilliant Blue, 25% isopropanol, 10% acetic acid.

3.1.28 Southern blot denaturing buffer - 0.5 M NaOH + 1.5 M NaCl.

3.1.29 Southern blot SSC wash buffer - 0.1x SSC + 0.1% SDS.

3.1.30 Southern blot transfer buffer - 20x SSC stock*: 3 M NaCl, 0.3 M N_3citrate*$2H_2O$. pH adjusted to 7.0 (with 1 M HCl).

3.1.31 TBE electrophoresis buffer - A 10x TBE stock was made and diluted as necessary, 109 g of Tris base, 55 g Boric acid and 4.65 g EDTA was diluted in 1 litre dH_2O and pH was adjusted to 8.3. The final concentration of TBE buffer was 89 mM Tris, 89 mM boric acid, 1.25 mM EDTA (pH 8.3).

3.1.32 TBS buffer - 140 mM NaCl, 25 mM Tris, 2.7 mM KCl, (pH 7.4).

3.1.33 TBS-T buffer - 140 mM NaCl, 25 mM Tris, 2.7 mM KCl, 0.05% Tween-20 (pH 7.4).

3.1.34 TE buffer* - 1 mM EDTA, 10 mM Tris (pH 7.5).

3.1.35 TSB solution - LB (pH 6.1) (adjusted with HCl), 10% PEG-3350, 5% DMSO, 10 mM $MgCl_2$, 10 mM $MgSO_4$.

3.1.36 2xTY* - 16 g/l Bacto-tryptone (Difco, UK), 10 g/l Bacto-yeast extract (Difco, UK), 5 g/l NaCl. Autoclaved.

3.1.37 Tris-glycine Transfer Buffer for western blot - 40 ml of 25x transfer buffer stock [12 mM Tris-base, 96 mM Glycine (pH 8.3)].

3.1.38 YE5S*: Yeast Extract with supplements - 5 g/l yeast extract, 30 g/l glucose, supplemented with 225 mg/l of adenine, histidine, leucine, uracil and lysine hydrochloride. Solid YE5S media was obtained by adding 2% Difco Bacto Agar to the solution. YES + 5 mg/l phloxin B (Sigma, UK) was used for investigating the ploidy of the cells. Phloxin B was added from a 5 g/l stock solution in sterile distilled water, when the medium had cooled below 60 ºC.

3.1.39 Western blot blocking solution - 4% milk powder in TBS-0.05% Tween 20.

3.1.40 Western blot washing buffer - TBS-0.05% Tween 20.

3.2 Methods.

3.2.1 Molecular biology based methods.

3.2.1.1 Bacterial transformation.

Conventional *E. coli* chemical transformation was carried out as follows: 2-3 μl of plasmid was added to 100 μl competent DH5α strain or Stbl cells (Invitrogen Ltd, UK) and mixed gently. Bacteria were incubated on ice for 20-30 minutes and heat shocked at 42 ℃ for 90 seconds. Bacteria were then kept on ice for 1-3 min. 1 ml of LB medium was added to bacteria and incubated 30-50 min at 37 ℃. 50 μl of cell solution was spread streaked on LB plates with appropriate antibiotics and grown at 37 ℃.

For TOP10 strain transformation, 3 μl DNA was added to a vial of TOP10 cells (Invitrogen Ltd, UK), incubated for 20 min on ice and then heat shocked at 42 ℃ for 30 seconds and placed on ice for 2 min. The cells were incubated in 200 μl SOC media at 37 ℃ for 50 min and then plated on LB plates with appropriate antibiotics and grown at 37 ℃.

Transformation of MAX Efficiency Stbl2 Competent Cells (Invitrogen Ltd, UK) was carried out according to manufacturer's instructions. Briefly, 3 μl DNA was added to 100 μl of competent cells, incubated on ice for 30 min, heat shocked at 42 ℃ for 25 seconds and placed on ice for 2 min. The cells were incubated in 900 μl SOC medium at 30 ℃ for 60 min, plated on LB plates with appropriate antibiotics and incubated at 30 ℃.

3.2.1.2 DNA digestion with restriction enzymes.

Samples were prepared in a total reaction volume of 30 μl by mixing 1 μg of DNA with 0.3 μl 100 x BSA, 10 U restriction enzymes and 3 μl of 10 x buffer compatible for enzymes that were being used. The sample was then vortexed briefly, centrifuged for 3 seconds at 13,000 rpm and incubated at 37 ℃ for 3

hours (if the samples were incubated overnight then 5 U of restriction enzyme was used). After incubation the enzymes were heat inactivated at 65 ℃ for 20 minutes. All enzymes were from NEB, USA.

3.2.1.3 Phosphorylation of the 5' termini of DNA.

Approximately 500 pmol of DNA was mixed with 2 μl of 10 x T4 ligase (NEB, USA) buffer, 10 U of T4 Polynucleotide Kinase (NEB, USA) and dH_2O to give a 20 μl final volume. After incubating the reaction for 1 hour at 37 ℃, the enzyme was heat inactivated by incubating the sample at 65 ℃ for 20 minutes or DNA purified by gel extraction (see 3.2.1.8).

3.2.1.4 De-phosphorylation of the 5' termini of DNA.

Approximately 500 ng of DNA was mixed with 2 μl of 10 x NEBuffer, 10 U of Calf Intestinal Alkaline Phosphatase (CIP) (NEB, USA) and dH_2O to give a 20 μl final volume. The sample was then incubated for 1 hour at 37 ℃ and the DNA purified by PCR purification (see 3.2.1.15) or gel extraction.

3.2.1.5 DNA ligation.

Ligation reaction was performed in 20 μl by mixing 2 μl 10 x T4 DNA ligase buffer (containing 10 mM ATP), approximately 100 ng of digested backbone and DNA insertion fragment in a molar excess of 3-10 times compared to the backbone (for two insert ligations into the backbone 4 times molar excess was used for inserts compared to the backbone). 400 U of T4 DNA ligase (NEB, USA) was added to the reaction, sample was mixed and incubated at 16 ℃ overnight.

3.2.1.6 Colony PCR.

A single bacterial colony was suspended in 50 μl LB medium. The PCR reaction was set up as follows: 1 μl of bacterial suspension in LB was added to a 29 μl PCR reaction that included the primers at 1.34 μM, 3μl 10x Taq buffer, 1.5 mM $MgCl_2$, dNTP's each at 0.125 mM and 2.5 U of Platinum Taq DNA polymerase

(Stratagen, UK). The reaction was initially denatured at 95 ℃ for 2 min, followed by 36 cycles of denaturation at 94 ℃ for 30 seconds, annealing at 58 ℃ for 1 min, extension at 72 ℃ for 2 min before a final extension at 72 ℃ for 8 min to generate the PCR product. The acquired PCR product was run on a 1% or 1.5% agarose gel for verification purposes.

3.2.1.7 Agarose gel electrophoresis.

Agarose was dissolved in TBE buffer, allowed to cool down to approximately 60 ℃, and ethidium bromide was added to 10 μg/ml. The gel was run in 1 x TBE buffer with 0.5 μg of DNA ladder (1 kb DNA ladder or λ phage ladder; Invitrogen, UK). Analysis was carried out using Gel Doc 2000 camera and Gel Doc software (Bio-Rad, UK) by exposing the gel to ultra violet light.

For DNA gel isolation, the agarose gel was prepared without ethidium bromide, and post run stained with DNA SYBR Green (1:10,000) in TBE for 15 minutes at RT. DNA was visualized with UV/BLUE converter plate, cut out with a scalpel and subjected to gel extraction.

3.2.1.8 Gel extraction of DNA.

DNA was then extracted from the gel slice using a QIAquick Gel Extraction kit (Qiagen, Germany) with a microcentrifuge according to the manufacturer's instructions. DNA was then eluted in either 50 μl or 30 μl of elution buffer EB, depending on the concentration of the original plasmid.

3.2.1.9 DNA sequencing.

0.5 μl ready Reaction Premix (ABI, USA) was mixed with 4 μl of ABI sequencing buffer (5x) and 3.5 μl ddH$_2$O per sequencing reaction and kept on ice. This was mixed with 12 μl of room temperature ddH$_2$O containing 3.2 pmol of primers and 150 ng of plasmid DNA. The reaction was initially denatured at 96 ℃ for 1 min, followed by 25 cycles of denaturation at 96 ℃ for 30 seconds, primer annealing at 50 ℃ for 20 seconds and extension at 60 ℃ for 4 min before

a final extension at 60 ℃ for 7 min. The reaction tubes or plate was briefly centrifuged and 5 µl of 125 mM EDTA was added. Thereafter, 60 µl of 100% ethanol was mixed with the reaction and left for 15 min at RT. Samples were centrifuged at 3,250 rpm for 30 min at RT or at 13,000 rpm in a benchtop centrifuge. Samples were inverted to remove the supernatant. 70 µl of 70% ethanol was added and centrifuged at 13,000 rpm for 15 min. The supernatant was removed and DNA dried at RT for 30 min.

3.2.1.10 Mini-preparation and sequencing in 96-well plates.

Individual bacterial colonies from LB plates with antibiotics were picked into individual wells of the 96-well plates containing 1.2 ml of LB medium. The plates were covered with an adhesive sealing film (ThermalSeal, Sigma, UK) with aeration holes and incubated overnight at 37 ℃ with constant rocking at 1300 rpm. The bacteria were then pelleted at 3,000 rpm for 15 min, the supernatant decanted and plates stored at -80 ℃ until mini-preparation. Mini-preparation and sequencing in 96-well plates were performed at the LRI sequencing facility.

3.2.1.11 Mini-preparation of plasmid DNA.

3 ml of LB medium with ampicillin (50 mg/l) or kanamycin (50 mg/l) was placed in a 5 ml polystyrene round-bottom tube (BD Falcon^TM) and inoculated with a single transformed E. coli colony. The culture was incubated overnight, shaking (225 rpm) at 37 ℃. 2 ml bacteria were harvested at 10,000 rpm for 3 minutes in a microcentrifuge. Isolation of plasmid DNA was then carried out using QIAprep Spin Miniprep Kit (Qiagen, Germany) according to the manufacturer's instructions. DNA was eluted with 50 µl Buffer EB (10 mM Tris-Cl, pH 8.5) or water, incubating it for 1 minute before centrifuging for 1 minute at 13,000 rpm. Plasmid DNA was stored at -20 ℃.

3.2.1.12 Midi-preparation of plasmid DNA.

Single bacterial colony was inoculated into 50 ml LB medium containing antibiotics (50 mg/l) and incubated for 24 hours at 37 ℃ with shaking (225 rpm). Isolation of plasmid DNA was then carried out using PureYield Plasmid Midiprep System (Promega, UK) according to the manufacturer's instructions. The DNA was eluted with 600 μl of nuclease-free water by adding it to the column and centrifuging it for 5 min at 1,600 rpm. Plasmid DNA was stored at -20 ℃.

3.2.1.13 Maxi-preparation of plasmid DNA.

A single bacterial colony was inoculated as a starter culture into 3 ml LB medium containing antibiotics (50 mg/l). The starter culture was incubated for 8 hours at 37 ℃ with shaking (225 rpm) and then used for inoculating 100 ml medium. Bacteria were grown at 37 ℃ for 12-16 hours with shaking (225 rpm). Bacterial cells were harvested by centrifugation at 6,000 x g in a Sorval GSA rotor for 15 min at 4 ℃ and supernatant removed by inverting the centrifuge tube. Isolation of plasmid DNA was then carried out using QIAGEN Plasmid Maxi Kit (Qiagen, Germany) according to the manufacturer's instructions. DNA pellet was air-dried for 10 min and re-dissolved in 0.5 ml TE buffer, pH 8.0. Plasmid DNA was stored at -20 ℃.

3.2.1.14 General PCR.

The general PCR procedure was as follows: 1 μl of DNA was added to a 49 μl PCR reaction that included primers at 0.5 μM, 5 μl 10x Pfu buffer, 5 μl 10x PCR enhance, dNTP's each at 0.4 mM and 2.5 U of Pfu Turbo R°DNA polymerase (Stratagene, UK). The reaction was initially at 94 ℃ for 2 min, followed by 9 cycles of 94 ℃ for 30 seconds, decreasing "touchdown" 63 ℃ by 1 ℃ every cycle and at 72 ℃ for 2 min. Then, 30 cycles at 94 ℃ for 30 seconds, at 55 ℃ for 30 seconds and at 72 ℃ for 3 min, was applied. PCR was carried out using an Eppendorf Mastercycler gradient PCR machine.

3.2.1.15 PCR purification.

For purifying PCR products from enzyme, nucleotides and primers, QIAquick PCR purification kit (Qiagen, Germany) was used according to manufacturer's instructions. 50 µl of Buffer EB or water was used to elute DNA. Alternatively, 30 µl of Buffer EB or water was used to elute DNA in order to increase the concentration of the DNA.

3.2.2 Methods related to mammalian or chicken cell experiments.

3.2.2.1 Reagents.

25 µg/ml lipopolysaccharide (LPS, Sigma, UK), 50 ng/ml mouse IL4 (R&D Systems Europe Ltd, UK) and 20 ng/ml human transforming growth factor (TGF) β1 (R&D Systems Europe Ltd, UK). The concentrations of estrogen [17-β-estradiol (Sigma, UK)], progesterone (Sigma, UK) are indicated in figure legends. Tamoxifen (Sigma, UK) was used at 50 nM unless stated; the final concentration for cycloheximide (Sigma, UK) was 10 µg/ml, 20 nM for actinomycin D (CalBioChem, USA), 4 µg/ml for α-amanitin (CalBioChem, USA) and 10 µM for MG-132 (Sigma, UK). Hormone stocks were prepared in DMSO at a concentration of 100 mM, this was then diluted in DMSO to give 1000 x stock solutions, and final dilutions were made in media (final DMSO concentration was less than 0.1 %). Hormone stocks were kept at -80 ºC, other reagents at either -20 ºC or 4 ºC according to manufacturer's guidance.

3.2.2.2 Mammalian cells lines.

Human cell lines Jurkat, SiHa, T47D, Ramos HS13 (derivate of Ramos), JAR and mouse B-cells and mouse B-cell lines were cultivated in RPMI 1640+GlutaMax medium (Invitrogen Ltd, UK), cell lines MCF7, JAMA2, HeLa and HepG2 in E4 medium and PC3 in HAMS F12 medium. Media was

supplemented with 10% FCS, 100 U / ml penicillin and 100 µg / ml streptomycin (Invitrogen Ltd, UK) and cells grown at 37 ℃ with 5% CO_2.

3.2.2.3 Chicken DT40 cell line.

Chicken B-cell derived DT40 cell line was cultured at 39 ℃ with 10% CO_2. Between 5 x 10^5 and 2 x 10^6 cells / ml were maintained in RPMI 1640+Glutamax media with heat inactivated (30 min at 55 ℃) 10% fetal calf serum, heat inactivated (30 min at 55 ℃) 1% chicken serum, 50 µM β-mercapthoethanol (Sigma, UK) and 10 U penicillin / streptomycin.

3.2.2.4 Mouse tissue.

Mouse tissue samples were derived from freshly sacrificed 8-12 week old female unplugged BALB/c mice. Mouse tissue was dissected (from 2 animals) and passed through cell strainers. Cells were incubated for 6 h in hormone-depleted media prior to estrogen treatment for 4 h.

3.2.2.5 Thawing cells.

A cryotube with cells was thawn and cells immediately transferred to a 15 ml falcon tube with 13 ml warm culture medium in it. Cells were centrifuged at 1,200 rpm for 8 min and the supernatant aspirated, leaving 100 µl of media on the pellet. The cells were resuspended by tapping the tube gently and transferred to a small sized flask with 6 ml culture medium.

3.2.2.6 Freezing cells.

Cell were grown to approximately 10^6 cells / ml. 5 x 10^6 cells were aliquoted into 15 ml falcon tubes, centrifuged at 1000 rpm for 5-7 minutes and the supernatant aspirated. Cells were gently resuspended in 1 ml ice cold freezing medium (10% DMSO in FCS) and transferred into a cryotube. The cells were placed into a NALGENE Cryo Freezing Container containing isopropanol, and stored at -80 ℃ for 72 hours before transferring the vials into liquid nitrogen.

3.2.2.7 Culturing cells in suspension.

Cells were monitored on a daily basis. The requirement for feeding the cells was estimated according to cell density, cell morphology and the color of the culture medium from red to orange/yellow. Cells were counted by using the haemocytometer.

Culture media was warmed up at 37 ºC prior to adding it to the cells. For chicken DT40 cells, 2.5 x volumes of media were added each day. For mammalian cells lines in cell suspension, approximately 2 x volumes of media were added each day. For adherent cells, the culture media was aspirated, cells rinsed with PBS and cells detached from the surface by trypsin treatment. Cells were then counted by haemocytometer and transferred to a new flask with fresh culture medium.

3.2.2.8 Hormone depletion from cells.

Cells were treated for 24 - 72 h (see 3.2.2.25) in hormone-depleted serum: Opti-MeM Reduced Serum Medium (Invitrogen Ltd, UK) supplemented with 10% Charcoal Stripped Fetal Bovine Serum (Invitrogen Ltd, UK), prior to hormone treatments. Opti-MeM Reduced Serum Medium was supplemented with: Charcoal Stripped Fetal Bovine Serum (Invitrogen Ltd, UK); Non-essential amino acids (final concentration and Sigma, UK catalog number): L-Alanine (8.9 mg/ml - A7469); L-Asparagine (15.0 mg/ml - A4159); L-Aspartic Acid (13.3 mg/ml - A4534); L-Glutamic Acid (14.7 mg/ml - G8415); Glycine (7.5 mg/ml - G7126); Proline (11.5 mg/ml - P5607; L-Serine (10.5 mg/ml - S2600).

3.2.2.9 Isolation of mouse splenic B-cells.

Mouse splenic B-cells were isolated with Mouse B-cell Negative Isolation Kit (Dynal Biotech, Norway) according to manufacturer's instructions. Mouse spleens were transferred to a tube containing 30 ml of RPMI 1640+Glutamax with 10% FCS preheated to 18-25 ºC. Spleens were sterilized by placing it twice into 70 % ethanol for 1 second and then into media. Tissues were then transferred to a wet 70 μm cell strainer on top of a 50 ml tube. Cells were macerated through the filter using the back of a syringe plunger. After this, the filter was rinsed

twice with 5 ml RPMI 1640 10% FCS solution. The tube was filled with cold RPMI 1640 10% FCS and centrifuged at 800 rpm at 4 ℃ for 10 minutes. Supernatant was discarded and cells resuspended in 5 ml room temperature RPMI 1640 1% FCS. The cells suspension was carefully pipetted on 7 ml Lympholyte M (Cedarlane Laboratories Limited, Canada) in a 15 ml tube. The sample was then centrifuged at 800 rpm for 20-30 minutes to remove erythrocytes. The interphase with B-cells was collected into a fresh tube. After filling the tube with RPMI 1640 1% FCS, cells were centrifuged at 800 rpm for 10 minutes and the supernatant removed, leaving 2-3 ml of media on the cells. 120 U of DNAse I (Roche, UK) was added per 1 ml cell suspension, mixed gently and incubated on a rotator for 15 minutes at RT. The suspension was filtered through a 70 μm cell strainer and counted for leucocytes. Thereafter, tube was filled with cold RPMI 1640 1% FCS and centrifuged at 800 rpm at 4 ℃ for 10 minutes. Leucocytes were suspended in PBS with 0.1% BSA (1 x 10^7 per 100 μl) and transferred to a 15 ml tube. 20 μl heat inactivated FCS per 1 x 10^7 leucocytes was added followed by the addition of 20 μl of Antibody Mix per 1 x 10^7 cells. The mix was incubated for 20 min at 4 ℃. Cells were then washed in 2ml PBS 0.1% BSA per 1 x 10^7 leucocytes and centrifuged at 800 rpm at 4 ℃ for 10 minutes. Supernatant was removed and cells resuspended in 800 μl PBS 0.1% BSA per 1 x 10^7 leucocytes. Mouse Depletion Dynabeads were washed in PBS 0.1% BSA before adding 200 μl per 1 x 10^7 leucocytes. Cells were incubated for 15 min at RT with gentle rotation. Bead-captured cells were resuspended and 1 ml PBS 0.1% BSA added per 1 x 10^7 leucocytes. The tube was placed in the magnet for approximately 2 minutes before transferring the supernatant with isolated mouse B-cells into a new tube. Cells were diluted in appropriate volume of RPMI 1640 with 10% FCS and grown at 37 ℃ with 5% CO_2.

3.2.2.10 Total RNA extraction from cell lines and mouse tissue derived cells.

The cells were shaken gently, then approximately 2 x 10^6 cells harvested and centrifuged in a Beckman Coulter Allegra^{TM} Centrifuge at 1,500 rpm for 7 min at RT. The supernatant was removed and the pellets frozen at -80 ℃ for at least 30 min to aid homogenization. The cell pellet was resuspended in 500 μl RNAWIZ buffer (Ambion, UK), homogenate incubated at RT for 5 min before the addition

of 100 μl of chloroform, vortexed for 20 seconds, incubated at RT for 10 min, centrifuged at 13,000 rpm in a benchtop centrifuge at 4 ºC for 15 min, and the upper aqueous phase was transferred to a new eppendorf tube. 250 μl of RNase free H_2O was added to the aqueous phase, mixed, followed by addition of 500 μl of isopropanol, mixed thoroughly, incubated at RT for 10 min, centrifuged at 13,000 rpm at 4 ºC for 15 min, and the supernatant was removed. The pellet washed with 500 μl of cold 75% ethanol, centrifuged at 13,000 rpm at 4 ºC for 5 min, the supernatant discarded and the pellet air-dried for 10-15 min before being resuspended in 20 μl RNase free H_2O. All tubes were RNase free and RNase free filtered pipette tips were used at all times when preparing and handling RNA. RNA samples were stored at -80 ºC.

Alternatively RNA was extracted using the RNeasy Mini kit (Qiagen, Germany) according to manufacturer's instructions. Where necessary, homogenisation of the sample was aided by passing the lysate at least 5 times through a 20-gauge needle fitted to an RNase-free suringe. RNA was eluted with 30 μl of RNase-free water by centrifuging the column at 13,000 rpm for 1 minute.

3.2.2.11 Synthesis of cDNA.

The genomic DNA was removed from total RNA samples prior to cDNA synthesis by using Turbo DNA-free Kit (Ambion, USA) according to the manufacturer's protocol. 2.5 μg of total RNA was denatured at 70 ºC for 2 min and then placed on ice, before adding 5x first strand buffer, DTT to a final concentration of 5 mM, dNTPs to a final concentration of 0.4 mM each, 0.1 μg oligo dT primer (Invitrogen Ltd, UK) and/or 2.5 μg $d(N)_{10}$ random primers (Promega, UK), 8 U RNase OUT (Invitrogen Ltd, UK), 200 U Superscript III reverse transcriptase (Invitrogen Ltd, UK) and the appropriate quantity of RNase free H_2O to a 20 ul reaction. This was incubated at 37 ºC for 10 min, 50 ºC for 40 min and then 55 ºC for 10 min. 80 ul of sterile H_2O was added to the cDNA before incubating at 100 ºC for 2 min.

3.2.2.12 Quantitative real time PCR (qRT-PCR).

The primers for real time PCR were designed by PrimerExpress software, using 60 ºC as the Tm and 150 bp as the length of the PCR product. Real time PCR analysis was carried out using the Abi 7000 Sequence Detection System (Applied Biosystems, UK) qRT-PCR machine and Abi software. Tissue or cell line hormone responsiveness was determined by qRT-PCR of GREB1 (gene regulated by breast cancer 1), a known estrogen-responsive gene. The real time PCR reaction was as follows: 10 µl of SYBRGreen PCR (Qiagen, Germany) reaction mixture was added to 9 µl of RNase free water and primers with a final concentration of 0.5 µM. 1 µl of cDNA was added to the PCR reaction. First, reaction was incubated at 50 ºC for 2 min, then at 95 ºC for 15 min followed by 40 cycles of melting at 95 ºC for 30 seconds, primer annealing at 56 ºC for 30 seconds and extention at 72 ºC. Real time PCR primers used in this study are listed in Table 1.

Primers SP3029 and SP3030 were used for amplifying mouse AID; SP3033 and SP3034 for mouse GAPDH; SP3065 and SP3066 for mouse GREB1; SP3067 and SP3068 for mouse c-myc. Real time PCR primers used for human AID were SP3055 and SP3056; SP3063 and SP3064 for human GAPDH. For amplification of pre-mRNA of mouse AID the following primers were used: SP3069 and SP3070 (for amplification of the region 75-226bp), SP3071 and SP3072 (2168-2318bp), SP3073 and SP3074 (4185-4335bp), SP3075 and SP3076 (6293-6443bp), SP3077 and SP3078 (7539-7689bp), SP3079 and SP3080 (9915-10065bp). Human AID was amplified by using SP3055 and SP3056; human GAPDH with SP3063 and SP3064. For amplifying AID cDNA from chicken derived DT40 cells, oligos SP3081 and SP3082 were used; for amplifying chicken GADPH, SP3083 and SP3084 were used.

Table 1. Primers used for gene expression analysis by real time PCR.

Gene	Species	Oligo	Direction	Sequence
Aid	mouse	SP3029	forward	5'-AACCCAATTTTCAGATCGCG-3'
Aid	mouse	SP3030	reverse	5'-AGCGGTTCCTGGCTATGATAAC-3'
Gapdh	mouse	SP3033	forward	5'-GCACAGTCAAGGCCGAGAAT-3'
Gapdh	mouse	SP3034	reverse	5'-GCCTTCTCCATGGTGGTGAA-3'

Greb1	mouse	SP3065	forward	5'-TCCGAGTTCAGAGGTCGGC-3'
Greb1	mouse	SP3066	reverse	5'-GTCCTACCTGTTGAGCTCCCACT-3'
Aid A	mouse	SP3069	forward	5'-GTCACGCTGGAGACCGATATG-3'
Aid A	mouse	SP3070	reverse	5'-AAAGACCTGAGCAGAGGGTGG-3'
Aid B	mouse	SP3071	forward	5'-TTCCCATCTAGGTAACACAGGA AGT-3'
Aid B	mouse	SP3072	reverse	5'-TCACCACCACAGTACAGGTA AACTC-3'
Aid C	mouse	SP3073	forward	5'-AATAAAA TCAACAAAACTGACCCAGC-3'
Aid C	mouse	SP3074	reverse	5'-CGAGGGATCAAACTCAAGACATC-3'
Aid D	mouse	SP3075	forward	5'-GAGGCCAATGACCGACCAC-3'
Aid D	mouse	SP3076	reverse	5'-CCATGGAAGCCAATCTGCA-3'
Aid E	mouse	SP3077	forward	5'-GCCACTCAGAGTGAGTGTCAGC-3'
Aid E	mouse	SP3078	reverse	5'-GAGGTCGGAGAATTGCAAGTTG-3'
Aid F	mouse	SP3087	forward	5'-TCTTGCCTCCTCTTCGCTCA-3'
Aid F	mouse	SP3088	reverse	5'-GTCTAAACGAAGTGGGTGGCTC-3'
Aid G	mouse	SP3079	forward	5'-GTAAGAGGGTGGCAAAATAGGGA-3'
Aid G	mouse	SP3080	reverse	5'-TTTACCAGAACCCAATTCTGGCT-3'
Apobec3	mouse	SP3153	forward	5'-TTCCACTGGAAGAGGCCCTT-3'
Apobec3	mouse	SP3154	reverse	5'-TCTCCAATCCTTTCCATGGC-3'
AID	human	SP3055	forward	5'-TCGGCGTGAGACCTACCTGT-3'
AID	human	SP3056	reverse	5'-GCCAGGGTCTAGGTCCCAGT-3'
GAPDH	human	SP3063	forward	5'-TCACCACCATGGAGAAGGCT-3'
GAPDH	human	SP3064	reverse	5'-CAGGAGGCATTGCTGATGATC-3'
APOBEC1	human	SP3151	forward	5'-TCCATGAGCTGCTCCATCAC-3'
APOBEC1	human	SP3152	reverse	5'-GCCGATTTTGTTGATCCATGT-3'
APOBEC2	human	SP3134	forward	5'-TGTGGAGCAAGAAGAGGGTGA-3'
APOBEC2	human	SP3135	reverse	5'-GCTGTCTCTTGGTGGCAGC-3'
APOBEC3A	human	SP3136	forward	5'-ACACACGTGAGACTGCGCAT-3'
APOBEC3A	human	SP3137	reverse	5'-GGTCCACAAAGGTGTCCCAG-3'
APOBEC3B	human	SP3138	forward	5'-GGAGCGGATGTATCGAGACAC-3'
APOBEC3B	human	SP3139	reverse	5'-CACCTGGCCTCGAAAGACC-3'
APOBEC3C	human	SP3140	forward	5'-CCAACGATCGGAACGAAACT-3'
APOBEC3C	human	SP3141	reverse	5'-TCGCAGAACCAAGAGAGGAAG-3'
APOBEC3D	human	SP3142	forward	5'-TGGCACTGATTGCAACTGACA-3'
APOBEC3D	human	SP3143	reverse	5'-GGCATGAATGGCTGACCTTC-3'
APOBEC3F	human	SP3144	forward	5'-ATTCATGCCTTGGTACAAATTCG-3'
APOBEC3F	human	SP3145	reverse	5'-GCTTTCGTTCCGACCATAGG-3'

APOBEC3G	human	SP3146	forward	5'-AAGTGGAGGAAGCTGCATCG-3'
APOBEC3G	human	SP3147	reverse	5'-AGTAGTAGAGGCGGGCAACG-3'
APOBEC3H	human	SP3148	forward	5'-CCCGCCTGTACTACCACTGG-3'
APOBEC3H	human	SP3149	reverse	5'-GGGTTGAAGGAAAGCGGTTT-3'
AID	chicken	GB3503	forward	5'-GATGCCTTTAAAACTCTGG-3'
AID	chicken	GB3527	reverse	5'-ATATCCTCCTCGATGTC-3'
GAPDH	chicken	SP3083	forward	5'-GCACTGTCAAGGCTGAGAACG-3'
GAPDH	chicken	SP3084	reverse	5'-GCCTTCTCCATGGTGGTGAA-3'

3.2.2.13 sIgM fluctuation analysis.

The surface IgM (sIgM) expression was investigated in Ramos HS13 cells (Sale and Neuberger, 1998), which contains a stop codon in the lambda locus, with reversion mutations resulting in sIgM production. Ramos HS cells were stained for cell surface expression of IgM μ-chain [(anti-human IgM–FITC (Sigma, UK)] and analysed with MoFlo (Cytomation, USA) flow cytometer. Single IgM negative cells were sorted into flat-bottom 96-well plates containing 200 μl of media and left to grow. 100 μl of fresh media with appropriate hormones concentrations was substituted for old media every 48 hours. Cells were treated with hormones for approximately 20 doublings. Thereafter, cells were once again stained for surface expression of IgM μ-chain and analysed by flow cytometry.

3.2.2.14 Isolation of genomic DNA.

Genomic DNA was isolated from cells using PUREGENE products Genomic DNA Isolation (Gentra Systems, USA) according to manufacturer's recommendations. 50 μl of TE (pH 8.0) were added to DNA pellets and left to rehydrate overnight at RT. Samples were stored at -20 ℃.

3.2.2.15 PCR amplification of VH, C, CD95/Fas, Sγ3 loci from genomic DNA.

VH and C regions were cloned and sequenced as published (Sale and Neuberger, 1998). The amplification of VH and C region from Ramos HS genomic DNA was as follows: 50 ng of DNA was added to a 50 μl PCR reaction that included

the primers VH forward and VH reverse at 20 μM, 10x buffer, 10x PCR enhance, dNTP's each at 0.4 mM and 2.5 U of Pfu Turbo R°DNA polymerase. The reaction was initially denatured at 94 ℃ for 1 min, followed by 9 cycles of denaturation at 94 ℃ for 30 seconds, decreasing annealing temperature from 68 ℃ by 1 ℃ every cycle, extension at 72 ℃ for 2 min. Thereafter, 35 cycles denaturation at 94 ℃ for 30 seconds, annealing at 63 ℃ and extension at 72 ℃ for 2 min was performed, before a final extension at 72 ℃ for 5 min to generate a 450 bp PCR product. Primer sequences used for VH amplification were VHF (5'-CCCCAAGCTTCCCAGGTGCAGCTACAGCAG-3') and VHR (5'-GCGGTACCTGAGGAGACGGTGACC-3'). For amplification of Ramos HS genomic C region in Ig locus the following primers were used: hCF (5'-CCCCAAGCTTCGGGAGTGCATCCGCCCCAACCCTT-3') and hCR (5'-CCCCGGTACCAGATGAGCTTGGACTTGCGG-3').

The human CD95/Fas locus was PCR amplified from isolated genomic DNA as published (Muschen et al., 2000). A 750-bp fragment encompassing the 5' region of the CD95/Fas gene was amplified using hFas1F (5'-ACCACCGGGGCTTTTCGTGA-3') as external forward, hFas2F (5'-TGAGCTCGTCTCTGATCTCG-3') as internal forward, hFas1R (5'-TATCTGTTCTGAAGGCTGCAG-3') as external reverse, and hFas2R (5'-CGGAGCGGACCTTTGGCT-3') as internal reverse primers in a nested PCR at an annealing temperature of 58 °C. The First PCR reaction contained 20 cycles of amplification, followed by 25 cycles during the second PCR reaction.

Genomic DNA was amplified and 750 bp of the Sγ3 region sequenced, using previously published (Xue et al., 2006) forward primer g3.2F (5'-gcgaattcTTGCAACTCC-TAAGAGGAAAGATCCC-3') and reverse primer g3.2R (5'-gcggatcCAGCCTGGTCC-CTACACTCCTAACAAC-3').

3.2.2.16 c-myc/IgH translocation.

c-myc/IgH translocations were detected by PCR as previously described (Kovalchuk et al., 1997; Ramiro et al., 2006), and carried out in collaboration with Isora V. Sernández (CNIO, Spain). Briefly, DNA was isolated from p53[+/-] B

cells after 72 hours of LPS stimulation in the presence or absence of 50 nM estrogen. Two rounds of PCR (primers listed in Table 2) were performed using Expand Long Template PCR system (Roche, USA) with primers MycIg1A and primers MycIg1B in the first round, and MycIg2A and MycIg2B in the second round. PCR products were separated on agarose gels, transferred to nylon membranes and probed with γ-$[P^{32}]$-ATP labelled oligonucleotides IgH probe and c-myc probe (Table 2). P values were calculated using two-tailed unpaired t test.

Table 2. Primers and probes used in c-myc/IgH tranlocation study.

MycIg1A	5'-TGAGGACCAGAGAGGGATAAAAGAGAA-3'
MycIg1B	5'-GGGGAGGGGGTGTCAAATAATAAGA-3'
MycIg2A	5'-CACCCTGCTATTTCCTTGTTGCTAC-3'
MycIg2B	5'-GACACCTCCCTTCTACACTCTAAACCG-3'
IgH probe	5'-CCTGGTATACAGGACGAAACTGCAGCAG-3'
c-myc probe	5'-GCAGCGATTCAGCACTGGGTGCAGG-3'

3.2.2.17 TOPO cloning of PCR products.

PCR fragments were cloned into TOPO vector using Zero Blunt TOPO PCR Cloning Kit (Invitrogen, UK) according to manufacturer's guidance. Cells were plated on LB plates with kanamycin and grown at 37 ºC overnight.

3.2.2.18 Western blot.

5 x SDS lysis buffer was added to 1 x to protein samples and denatured at 90 ºC for 5 minutes. The samples were then left to cool down for 3 minutes, 4 x LDS buffer was added to 1 x and the samples incubated at 70 ºC for 10 minutes. 4 μl of Prestained SDS-PAGE Standand Low Range Ladder (Bio-Rad, UK) or 0.5 μl of BenchMark Ladder (Invitrogen, UK) was electrophoresed simultaneously with the protein samples using either MOPS SDS Running Buffer (Invitrogen, UK) or MES SDS Running Buffer (Invitrogen, UK) at 200V for 50 minutes at RT. The separated proteins were transferred onto Immobilon-P polyvinylidene fluoride (PVDF) membrane (Millipore, UK) in transfer buffer in a Novex MiniGel transfer apparatus (Invitrogen, UK) at 30 V for 150 minutes on ice. Membranes

were blocked with 4% milk in TBS-0.05% Tween 20 (blocking solution) for 30 min at RT or overnight at 4 ºC, probed with primary antibody (in blocking solution) for 1 h at RT followed by washing the membrane with TBS-0.05% Tween 20 three times for 5 minutes at RT. The membrane was then incubated with HRP-conjugated secondary antibody in blocking solution for 1 h at RT, washed 3 times with TBS-0.05% Tween 20 for 5 minutes at RT, and bands detected with ECL Plus Western Blotting Detection System (GE Healthcare, UK) according to manufacturer's recommendations. Primary and secondary antibodies used for western blotting are listed in Table 3.

Quantitative western blot was carried out using the LI-COR system, according to manufacturer's instructions. Potein was visualized by IR Dye 680 and analysed with LI-COR Odyssey Infrared Reader.

Table 3. Antibodies used for western blot.

Antibody	Concentration / dilution
Mouse monoclonal anti-FLAG M2 (Sigma, F3165)	0.25 µg/ml
Mouse monoclonal anti-Myc 9E10 (LIF Monoclonal Antibody Service)	0.5 µg/ml
Mouse monoclonal anti-RACK1 (BD Biosciences, 610177)	0.1 µg/ml
Rabbit polyclonal anti-PARP1 (Alexis Biochemicals, ALX-210-219-R100)	1/10,000
Rabbit polyclonal anti-PAR (Trevigen, 4336)	1/1,000
Mouse monoclonal anti-Tubulin (LIF Monoclonal Antibody Service)	1/10,000
Goat anti-mouse Ig-HRP conjugate (DakoCytomation, P0449)	1/10,000
Goat anti-rabbit Ig-HRP conjugate (DakoCytomation, P0448)	1/2,000
Goat anti-mouse IR Dye 680 (LI-COR Biosciences, 926-32220)	1/15,000

3.2.2.19 Circle transcript detection by RT-PCR in mouse B-cells.

Isolated splenic B-cells were stimulated for up to 72 h with LPS + IL4 for inducing switching to IgG1 and IgE, LPS + TGF-β for switching to IgA and LPS for switching to IgG3. Hormones were added to the cells together with LPS and cytokines in fresh media. Primers (Table 4) that were used for detecting IgG1 (Reina-San-Martin et al., 2003), IgG3, IgA (Kinoshita et al., 2001) and IgE (Lumsden et al., 2004), have been described previously.

Table 4. Primers used for detecting circle transcripts.

Subclass	Orientation	Sequence
IgG1	Forward	5'-TCGAGAAGCCTGAGGAATGTG-3'
IgG1	reverse	5'-GAAGACATTTGGGAAGGACTGACT-3'
IgG3	forward	5'-TGGGCAAGTGGATCTGAACA-3'
IgG3	reverse	5'-AATGGTGCTGGGCAGGAAGT-3'
IgA	forward	5'-CCAGGCATGGTTGAGATAGAGATAG-3'
IgA	reverse	5'-AATGGTGCTGGGCAGGAAGT-3'
IgE	forward	5'-TTGGACTACTGGGGTCAAGG-3'
IgE	reverse	5'-CAGTGCCTTTACAGGGCTTC-3'

3.2.2.20 Promoter analysis.

Bioinformatic promoter analysis was carried out with the AliBaba2.1 transcription factor response element predicting program using Transfac 4.0 transcription factor binding site database as a source for constructing matrices. Human AID promoter fragments were analyzed using pE1BLuc plasmid backbone (kindly provided by Göran Akusjärvi, Uppsala University) containing a luciferase ORF and a minimal promoter region. Primers (Table 5) approximately every 500 base pairs from the human AID transcription start site were used to amplify AID promoter regions. The PCR procedure was as follows: 1 μl of genomic DNA was added to a 49 μl PCR reaction that included primers at 0.5 μM, 5 μl 10x Pfu buffer, 5 μl 10x PCR enhance, dNTP's each at 0.4 mM and 2.5 U of Pfu Turbo. The reaction was initially at 94 ºC for 2 min, followed by 9 cycles of 94 ºC for 30 seconds, (decreasing "touchdown" 63 ºC by 1 ºC every cycle) and at 72 ºC for 3 min. Then, 30 cycles at 94 ºC for 30 seconds, at 55 ºC for 30 seconds and at 72 ºC for 3 min, was applied. Products were cloned into pE1BLuc vector and 3 μg DNA were transfected into SiHa cells using Lipofectamine 2000 (Invitrogen, UK) according to manufacturer protocols. 24 h after transfection, cells were treated with indicated concentrations of hormones or TNF-α for 4 h and analyzed thereafter for luciferase signal by using the Dual Luciferase Reporter Assay System (Promega, UK) and Glomax luminometer (Promega, UK). The expression vector containing a dominant negative mutant for IκBα (S32A S36A - kindly provided by Felix Randow, MRC-LMB, UK), was cotransfected with pE1BLuc constructs where indicated. One 1μg CMV promoter

driven Renilla luciferase expression vector (Promega, UK) was included in all transfections and used as a control for transfection efficiency.

Table 5. Primer sequences for hAID promoter analysis.

Oligo name	Approximate position relative to transcription start	Direction	Sequence
SP3089	0bp	forward	5'-CTCGAGGCCAATGCACTGTCAGACTA-3'
SP3090	0bp	reverse	5'-CAGCTGGAAAAATCTCACTTCAATTAATGAT-GGTTC-3'
SP3092	-2000bp	forward	5'-CTCGAGGATGGTGTAAGCCACAACCA -3'
SP3093	-1500bp	forward	5'-CTCGAGCAAGAAGAGTAGGTAAGGCAG-3'
SP3094	-1000bp	forward	5'-CTCGAGATTTGAAAATCATCAAGGTATAGATG-3'
SP3095	-500bp	forward	5'-CTCGAGACTGAGTTCATTTGCTTAACTGCA-3'
SP3096	500bp	forward	5'-CAGCTGTTACAAAATTATTACGAAAATTAGCACT-ACC-3'
SP3099	2000bp	reverse	5'-CAGCTGTCTCTTTGAGGCCCAGTG-3'
SP3108	1500bp	forward	5'-CAGCTGTTTCAGGCTTGCAGGCTGACAG-3'
SP3101	pE1BLuc	forward	5'-ACATATTGTCGTTAGAACGCGGCTAC-3'
SP3102	pE1BLuc	reverse	5'-CCAACAGTACCGGAATGCCAAG-3'

3.2.2.21 Preparation of nuclear extracts.

For isolating nuclear extracts from 5×10^6 Ramos, the Pierce Nuclear Extraction Kit was used according to manufacturer's instructions. Supernatant (nuclear extract) fractions were stored at -80 ºC.

3.2.2.22 Electro-mobility shift assay (EMSA).

EMSAs were performed by standard protocol (Ausubel et al., 1987) with minor modifications. Complementary oligonucleotides containing either NFκB or putative estrogen receptor binding elements (ERE) were annealed to each other and end labeled with [γ-^{32}P]ATP by T4 polynucleotide kinase to a specific activity of ca. 300,000 to 500,000 cpm/ng. Seven μg of nuclear extracts were

incubated with the binding mixture containing 20 mM HEPES (pH 7.8), 50 mM KCl, 10% glycerol, 0.25 mM DTT (Invitrogen Ltd, UK), 0.1 mM EDTA, 0.55 µg of poly(dI-dC) (GE Healthcare, UK) and 0.25 pmol of labeled oligonucleotides for 20 min at RT in a total volume of 12 µl. Competition assays included 1-, 3- and 10-fold mass excess of unlabeled oligonucleotides. Antibodies to NFκB p65 (polyclonal - Santa Cruz, USA, sc-109), ERα (monoclonal, Abcam, UK, ab2746) or mouse λ (polyclonal, Southern Biotech, USA, 1060-01) were added 10 min after starting the reaction (0.2 - 1.8 µg / assay) and incubated for further 20 minutes at RT. Gels were pre-electrophoresed at 10 mA for 60 min. Samples were electrophoresed at 4 ˚C on a pre-run non-denaturing 4.5 % poly-acrylamide gels (30:1) in low-ionic-strength TBE buffer (10 mM Tris-borate, pH 7.5; 0.025 mM EDTA) at 20 mA / gel for 120 min. This was followed by autoradiography on imaging plates (FujiFilm, Japan) and analysis with a FLA-5000 Scanner (FujiFilm, Japan).

3.2.2.23 Chromatin immunoprecipitation (ChIP).

Approximately 1 x 10^7 cells were fixed with 1 % formaldehyde in the culture media for 10 min at 37 ºC, then quenched with 0.125 M glycine for 5 min at RT. Cells were pelleted and washed twice with cold PBS. The cell pellet was resuspended in 300 µl SDS lysis buffer [1% SDS, 10 mM EDTA, 50 mM Tris-HCl (pH 8.1)] and incubated on ice for 10 min. Cell lysates were sonicated for 6 min with a 30 sec on/off sonication cycle in 1.5 ml eppendorf tubes. Samples were centrifuged for 5 minutes at 13,000 rpm at 4 ºC and supernatants diluted to 1.5 ml with ChIP buffer [0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl (pH 8.1), 167 mM NaCl] - an aliquot was kept as input measurement. Samples were pre-cleared with 60 µl Protein G beads (ProtG - Roche, UK) and 1 µg salmon sperm DNA (Invitrogen Ltd, UK) for 45 min at 4 ºC. Anti-NFκB p65 antibody (Santa Cruz, USA, sc-109), rabbit anti-ERα HC-20 antibody (Santa Cruz, USA, sc-543x), mouse anti-PR-AB (monoclonal, Abcam, UK, ab2764) or goat anti-mouse lambda control antibody (see above) was added to the supernatant at a concentration of 2 µg / assay and incubated for 12 h at 4 ºC. Sixty µl of salmon sperm DNA (1 µg DNA / 20 µl ProtG) was added to the samples and incubated for 1 h while rotating. The ProtG/ antibody/ protein

complex was pelleted and washed with the following buffers: 1) low salt wash buffer [0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl (pH 8.1), 150 mM NaCl], 2) high salt wash buffer [0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl (pH 8.1), 500 mM NaCl], 3) LiCl wash buffer [0.25 M LiCl, 1% IGEPAL CA630, 1% deoxycholic acid, 1 mM EDTA, 10 mM Tris-HCl (pH 8.1)], TE buffer [10 mM Tris-HCl (pH 8.0), 1 mM EDTA], and twice with a standard TE buffer. The sample was eluted with 250 µl of elution buffer [1% SDS, 0.1 M NaHCO3 (pH 7.5)] and incubated for 15 min at RT with agitation. The beads were centrifuged and the elution was repeated once. Bound immunocomplexes were then reverse-crosslinked with 200 mM NaCl by incubating it at 65 ºC for 12 h. Ten mM EDTA (pH 8), 40 mM Tris-HCl (pH 6.5) and 20 µg of proteinase K was added to the sample and incubated for 1 h at 45 ºC. DNA was then extracted with phenol/chloroform and precipitated with 130 mM NaOAc (pH 5.5), 30 µg glycogen (Roche, UK) and 50% ethanol for 2 h at -20 ºC. The DNA pellet was washed with 70% ethanol, resuspended in 50 µl of water and subjected to qRT-PCR, using forward primer CHIPEF (5'-GCGTGAGCTCTCTCTTGCCT-3') and reverse primer CHIPER (5'-CACTGCTAATAAAGACATGCCTGAG-3') for amplifying -1189 to -1039 bp of AID promoter region.

The following qRT-PCR primers were used in anti-PR ChIP studies: forward primer CHIP1 and reverse primer CHIP2 for region I, forward primer CHIP3 and reverse primer CHIP4 for region II, forward primer CHIP5 and reverse primer CHIP6 for region III, forward primer CHIP7 and reverse primer CHIP8 for region IV, forward primer CHIP9 and reverse primer CHIP10 for region V (Table 6).

Table 6. Primer sequences used for anti-PR ChIP studies.

CHIP1	forward	5'-GCAGGAAGGCCAGTTGGGAA-3'
CHIP2	reverse	5'-GCTCCTCTCTTTCTTTTAAATTCCATGCAC-3'
CHIP3	forward	5'-GACAGGCTGGAGGTGCTCAG-3'
CHIP4	reverse	5'-GACAGGCTGGAGGTGCTCAG-3'
CHIP5	forward	5'-CCACATGTCATGGGAGGGA-3'
CHIP6	reverse	5'-GTGGTAGGCAAGAGAGAGCTCAC-3'

CHIP7	forward	5'-TCAGGCATGTCTTTATTAGCAGTGT-3'
CHIP8	reverse	5'-GATGCAATGCTAAGCAATGCAG-3'
CHIP9	forward	5'-GATCTTAGCCTGTTTTCCAAATTCAG-3'
CHIP10	reverse	5'-TTTTGCATGAGCCATCTTTTTTAG-3'

3.2.2.24 Lipofectamine transfection of SiHa cells.

Transfection of SiHa cells with Lipofectamine 2000 (Invitrogen, UK) was performed as suggested by the manufacturer. The day before transfection, cells were plated into 96-well plates in a density so that the next day, cells should have a confluency of 90-95%. The media used for Lipofection was Opti-MEM Reduced Serum Medium. Cells were grown in Opti-MEM supplemented with 10% Charcoal Treated Fetal Bovine Serum. Approximately 5 hours before transfection, fresh media with FCS but without antibiotics was added to the cells. Lipofectamine was diluted in Opti-MEM medium without serum and incubated for 5 min at RT. Plasmid DNA was diluted in Opti-MEM medium without serum and combined with Lipofectamine, followed by incubation at RT for 20 min. The medium in plates was substituted with serum free medium and the DNA-lipofectamine complexes were added to the wells. The cells were incubated for 6 hour at 37 ºC and the medium was once again substituted with medium containing 10% FCS. After incubation for 24 hours, fresh media with serum and antibiotics were added to the cells.

3.2.2.25 Time and concentration of hormones used during various experiments.

A list of different concentrations of estrogen and progesterone and the time of cell treatment for various experiments is indicated in Table 7.

Table 7. Treatment of cells with hormones.

Experiment	Cells	Depletion	Hormone	[nM]	h
AID mRNA	B-cells	24	Estrogen	0.1-1,000	1-8
			Progesterone	0.1-10,000	1-24
AID mRNA	Ramos	72	Estrogen	10-100	1-4
			Progesterone	200-2,000	1-4
AID mRNA	DT40	48	Estrogen	10-100	1-8
			Progesterone	200	1-8
AID pre-mRNA	B-cells	24	Estrogen	1-10	1-8
			Progesterone	200	1-24
AID protein	DT40	48	Estrogen	10-100	1-8
			Progesterone	200	1-8
AID EMSA	Ramos	72	Estrogen	10	4
AID ChIP	Ramos	72	Estrogen	1-10	4
			Progesterone	200	4
SHM	Ramos	72	Estrogen	10-100	3 wks
			Progesterone	200-2,000	3 wks
SHM	B-cells	24	Estrogen	10	6 days
			Progesterone	200	6 days
CSR (switch circle)	B-cells	24	Estrogen	1-10	48
			Progesterone	200	48
CSR FACS	B-cells	24	Progesterone	200	6 days
c-myc/Ig	B-cells	24	Estrogen	50	72
Promoter analysis	SiHa	24	Estrogen	0.1-1,000	4
			Progesterone	200	4
Mouse tissues	Various	6	Estrogen	1-10	4

3.2.3 Methods for generating endogenously tagged AID DT40 cell lines.

The generation of tagged AID constructs was undertaken in collaboration with Gudrun Bachmann.

3.2.3.1 Generating tags.

Primer sequences (Table 8) for the following tags were generated: 3FLAG (abbreviated FLAG), 3Myc (abbreviated Myc), 3FLAG-2TEV-3Myc (abbreviated FLAG-TEV-Myc) and eGFP (Figure 8).
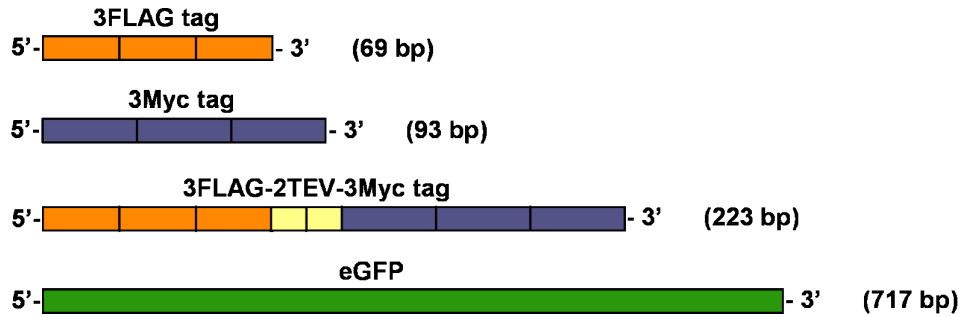
Figure 8. A schematic illustration of the tags that were attached to 3' end of chicken AID. The FLAG and Myc epitopes are in three repeats whereas the TEV protease cleavage site is in two repeats. The size of each individual tag is shown at the 3' end of the respective tag. Not drawn to scale.

3FLAG: Oligonucleotide F1 and F2 were hybridized and the sample subjected to electrophoresis on a 10% PAGE gel. The gel was stained with SYBR green, visualized with Gel Doc 2000 UV-illuminator and the FLAG fragment cut out of the gel. The fragment was melted into a 1.5% agarose gel before being subjected to electrophoresis and purification by gel extraction.

3Myc: Oligonucleotide M1 and M2 were hybridized and purified as described for 3FLAG tag.

3FLAG-2TEV-3Myc: Oligonucleotide FT1, FT2 and FT3 were separately 5'-end phosphorylated by T4 polynucleotide kinase (PNK). Thereafter, FT1, FT2 and FT3 were hybridized to FT4, FT5 and FT6, respectively, yielding fragment A, B and C. Fragment A and B were then ligated by T4 DNA ligase, yielding fragment AB. The AB fragment was purified as above. The purified AB fragment was 5'-end phosphorylated and ligated to fragment C, yielding fragment ABC. The ligation sample was subjected to electrophoresis and purification from the gel. The ABC fragment was then ligated to Spe I digested, calf intestinal alkaline phosphatase dephosphorylated, gel-purified plasmid containing 3Myc sequence. The FLAG-TEV-Myc fragment was amplified by Pfu Turbo°R DNA polymerase using forward primer FTM1, reverse primer FTM2 and the ligation sample as the template. The PCR parameters were as follows: denaturation at 94 ºC for 30 seconds, primer annealing at 53 ºC for 30 seconds and polymerization at 72 ºC

for 45 seconds, for 35 cycles. The PCR sample was purified by PCR Purification Kit and digested with Spe I and EcoR I. Thereafter, the sample was run on a 1.5% agarose gel, followed by purification from the gel.

eGFP: The eGFP open reading frame (ORF) was amplified from a pIRES-eGFP vector by using Pfu Turbo R°DNA polymerase, forward primer G1, reverse primer G2, and the following PCR parameters: denaturation at 94 ℃ for 30 seconds, primer annealing at 48 ℃ for 30 seconds, polymerization at 72 ℃ for 70 seconds, for 35 cycles. The sample was run on a 1% agarose gel, followed by purification from the gel. Thereafter, the fragment was double digested with Spe I and EcoR I and purified with the PCR Purification Kit.

Table 8. List of oligos used in generating tags.

Oligo name	Oligo sequence
F1	5'-CTAGTGACTACAAGGACCACGATGGCGACTACAAGGAT CACGACATCGATTACAAGGACGATGACGACAAGTG-3'
F2	5'-AATTCACTTGTCGTCATCGTCCTTGTAATCGATGTCGTGA TCCTTGTAGTCGCCATCGTGGTCCTTGTAGTCA-3'
M1	5'-CTAGTGAGCAGAAGCTGATCTCCGAGGAAGACCTGGAGCAG-AAGCTGATCTCCGAGGAAGACCTGGAGCAGAAGCTGATCTCCGA-GGAAGACCTGTG-3'
M2	5'-AATTCACAGGTCTTCCTCGGAGATCAGCTTCTGCTCCAGGTCTT-CCTCGGAGATCAGCTTCTGCTCCAGGTCTTCCTCGGAGATCAGC-TTCTGCTCA-3'
FT1	5'-TACAAGGACGATGACGAACAAGGATTACGATATCCCAACGAC-CGAA-3'
FT2	5'-CTTGTAATCGATGTCGTGATCCTTGTAGTCGCCATCGTGGTCCTT-3'
FT3	5'-CCTGTATTTTCAGGGCGAAAACCTGTATTTTCAGGGCC-3'
FT4	5'-TACAGGTTTTCGGTCGTTGGGATATCGTAATCCTTGTCGTCATCG-3'
FT5	5'-CTAGTGACTACAAGGACCACGATGGCGACTACAAGGATCACGACA-3'
FT6	5'-CTAGGGCCCTGAAAATACAGGTTTTCGCCCTGAAAA-3'
FTM1	5'-CAGGTCGACTCTAGAG-3'
FTM2	5'-CTAGGGCCCTGAAAATACA-3'
G1	5'-ATGGTGAGCAAGGGC-3'
G2	5'-CTTGTACAGCTCGTCC-3'

3.2.3.2 Generating constructs with tagged AID sequence.

A part of chicken AID gene (Genbank accession number XP_416483) was inserted into a pBluescript plasmid (Figure 9) containing a puromycin resistance gene flanked with *LoxP* sites. The vector was double digested with Spe I and EcoR I followed by CIP dephosphorylation, electrophoresis and purification from the gel.
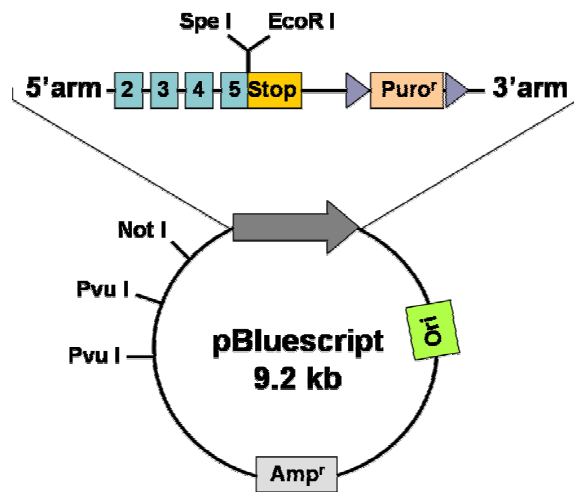
Figure 9. The plasmid containing the chicken AID gene. The pBluescript plasmid contains a part of the chicken AID gene and a puromycin (pink) selection cassette. *loxP* sites, depicted as blue triangles, are located at the 5'- and -3'-end of the selection cassette. 5'-arm is 1 kb of the intron between exon one and exon two in the chicken gene, whereas -3'arm stands for approximately 1 kb of the sequence that lies downstream of the AID gene. Exons are depicted as light blue, numbered squares. Stop codon is shown as a orange square. *ori* stands for origin of replication and Amp^R for a gene coding for β-lactamase.

Next, each tag was ligated into the linearized backbone and transformed into bacteria. Clones were cultured and subjected to plasmid purification. The purified DNA was verified by Stu I and Nsi I digestion and visualization on an agarose gel and also sequenced using the BigDyer Terminator v3.1 Cycle Sequencing Kit and primers listed in Table 9. Sequences were analyzed using Sequencer*TM* 4.5.

Table 9. **Primers for sequencing tagged AID constructs.** Sequences of HPLC purified primers used for sequencing tagged AID constructs. '*' denotes primers used for sequencing construct with eGFP tag.

Name	Sequence
S1	5'-ACCGTTACCTTAAAATACTGC-3'
S2	5'-GTCCCTGCTGCTTTAAC-3'
S3	5'-GGATATGATGATTTAGTGAGC-3'
S4	5'-GATGCCTTTAAAACTCTGG-3'
S5	5'-AATTCAGAGGAAGTCATCAG-3'
S6	5'-TACAAATGTGGTATGGCTGA-3'
S7	5'-ATCTCGGCGAACACC-3'
S8	5'-TCACACGCCAGAAGC-3'
S9	5'-TGCTCAGCAACTCGG-3'
S10	5'-GCGCGCTTCGCTTTT-3'
S11	5'-CTATGACAGGTTGAAACTAG-3'
S12	5'-TTGGGGTTATGTGAGTTC-3'
S13	5'-AAGGGACCCAATCATATCT-3'
S14	5'-ACTGTCCACAAAACCAGATA-3'
S15	5'-TCAAAAGAACTCACATAACCC-3'
S16	5'-GTCTAGTTTCAACCTGTCA-3'
S17	5'-TTTCCTTTTATGGCGAGG-3'
S18	5'-TCCCCGAGTTGCTGA-3'
S19	5'-TTCTCCCTCTCCAGC-3'
S20	5'-TGTTCGCCGAGATCG-3'
S21	5'-GGATCATAATCAGCCATAC-3'
S22	5'-TATTGCTGATGACTTCCTC-3'
S23	5'-CTAGTAAGTCCCAGAGTTT-3'
S24	5'-TCTCTTCTAGGCTCACTAA-3'
S25	5'-CTGAGGTACTGTTAAAGCA-3'
S26	5'-GCAGTATTTTAAGGTAACGG-3'
S29*	5'-AGGAGGACGGCAACAT-3'
S30*	5'-TGTAGTTGTACTCCAGCT-3'
m13F	5'-GTTTTCCCAGTCACGAC-3'
m13R	5'-GGAAACAGCTATGACCATG-3'

Clones transformed with the construct containing AID-FLAG-TEV-Myc, were picked and resuspended in 2xTY broth. The resuspension was used as a template in modified colony PCR, using forward primer 5'-CTGCCACTGTATGAAGT-3', reverse primer 5'-CAGGTCGACTCTAGAG-3'

and the following PCR parameters: denaturation at 94 ℃ for 30 seconds, primer annealing at 50 ℃ for 60 seconds, polymerisation at 72 ℃ for 45 seconds, for 30 cycles, were applied after initial denaturing and lysis step at 95 ℃ for six minutes. Positive clones were grown further and used for isolating plasmid and sequencing. A similar procedure was carried out for all constructs. Samples were sequenced for plasmids and digested with Not I, or Pvu I, followed by ethanol-precipitation and air drying of the DNA. The DNA pellet was then resuspended in sterile PBS buffer.

3.2.3.3 Stable transfection of DT40 cells.

Newly thawed DT40 (CL18⁻) cells were incubated in filter-sterilized culture medium (10% fetal calf serum, 1% chicken serum, 50-75 μM β-mercaptoethanol, 0.01% penicillin, 0.01% streptomycin in RPMI 1640+Glutamax media at 39 ℃ in 10% CO_2 for three days. The cells were maintained at a density of $5x10^5$-$1x10^6$ cells / ml by splitting the culture every day (described above in Cell culturing section). $1x10^7$ cells were then spun down at 90 x g for 5-10 minutes and resuspended in RPMI 1640+Glutamax medium (without FCS) to give a 500 μl final volume. 25-50 μg of Not I or Pvu I linearized plasmid containing tagged AID was added to the culture solution and transferred to a 4 mm electrode gap Gene Pulser Cuvette. The cuvette was placed on ice for 10 minutes and electoporated using the following parameters: voltage = 550 V, capacitance = 25 μF, resistance = ∞, followed by incubating the culture on ice for five minutes. The cells were transferred to a flask, containing filter-sterilized culture medium with serum and incubated for 8 hours at 39 ℃ in 10% CO_2. The culture was then dispensed into 96 well flat bottom plates and incubated overnight. Puromycin was thereafter added to each well, to a final concentration of 50 μM, and the plates incubated for 5-7 days. Clones were fed or splitted every day until reaching high density. The overall scheme for generating endogenously tagged AID DT40 cell lines is shown in Figure 10.
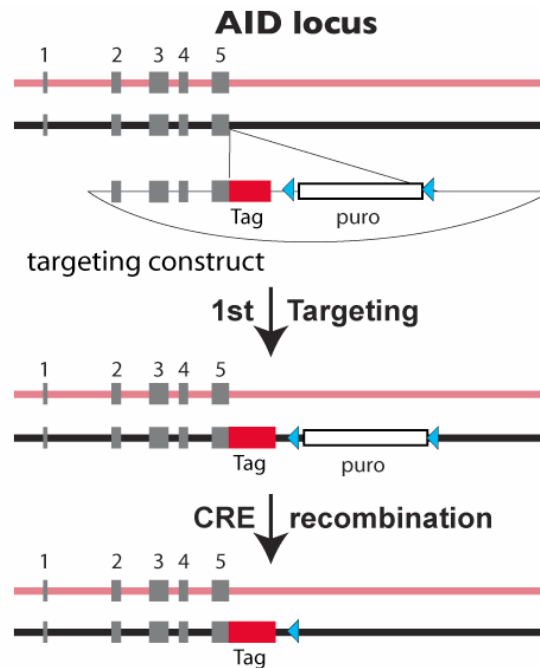
Figure 10. Endogenous tagging of AID in DT40. Schematic representation of the targeting of chicken AID locus with the tagged AID construct (AID-tag). A targeting construct that contains a part of chicken AID genomic locus, an in-frame tag (red rectangle) and a puromycin selection cassette (white rectangle) between two loxP sites (blue triangles) is targeted into chicken AID allele by homologous recombination. After confirming the targeting of the allele by southern blotting, puromycin cassette is removed by Cre recombinase.

3.2.3.4 Verifying stably transfected DT40 cell lines.

A southern blotting probe was prepared by PCR amplifying a part of the AID gene using Pfu Turbo°R DNA polymerase, forward primer Pa1 (5'-CTTATGACTGTGCCCGACATG-3'), reverse primer Pa2 (5'-CTGCTCCAGAGAGGACAC-3') and the generated FLAG-tagged plasmid as a template. The PCR parameters were the following: denaturation at 94 °C for 30 seconds, primers were annealed at 57 °C for 30 seconds, polymerization at 72 °C for 60 seconds, for 35 cycles. Genomic DNA from $6\text{-}9\text{x}10^6$ cells was purified using the PUREGENE DNA Isolation Kit followed by digestion with Mph1103I endonuclease or with Spe I. The digested DNA was subjected to electrophoresis on a 0.7% agarose gel. This was followed by incubating the gel in denaturation buffer (0.5 M NaOH, 1.5 M NaCl) for 45 min at RT with constant agitation. The DNA was then blotted to a Hybond-N+ membrane, the membrane washed in 2xSSC (30 mM trisodium citrate, 300 mM NaCl) and exposed to 0.12 joules of

UV light using an UV-crosslinker. The membrane was placed in a roller bottle, 65 ºC warm ExpressHyb solution was added to the bottle and incubated at 65 ºC for one hour. The probe was [α^{32}P] dCTP labelled using RediprimeTM II Random Primer Labelling System. Thereafter, the radioactive probe was mixed with 5 ml ExpressHyb solution, added to the roller bottle and incubated at 65 ºC for 2 hours. The membrane was washed in wash buffer (1.5 mM trisodium citrate, 15 M NaCl, 0.1% SDS) at 65 ºC, followed by exposing the membrane to a phosphorimager cassette. The radioactive signal was detected with a FLA 5000 phosphorimager.

3.2.3.5 Transient transfection of DT40 cells with Cre recombinase.

$3x10^6$ cells were centrifuged at 1000 rpm for 5-10 minutes, the medium was aspirated and cell pellet resuspended in 100 μl supplemented NucleofectorTM Solution from the cell line Nucleofector°R kit (Amaxa, UK). This was mixed with 4 μg of pCMV Cre plasmid, or 2 μg pmaxGFP plasmid. The sample was transferred to an Amaxa cuvette, electroporated using B023 NucleofectorTM program, the cells transferred to a 6 well plate containing culture medium and incubated for 30 hours. The transfection efficiency was then analyzed by GFP detection by flow cytometry. Next, the pCMV Cre transfected cells were dispensed into 96-well flat bottom plates, one clone per well, and incubated for 10 days. Each colony of clones was then divided into two wells and puromycin was added (final concentration 50 μM) to one of the duplicates, followed by incubation for 2-3 days. The wells that contained puromycin, were investigated for cell growth and survival by visual estimation.

3.2.3.6 Verifying puromycin sensitive clones by southern blot.

A southern blot probe was amplified from the FLAG-tagged plasmid with Pfu Turbo R° DNA polymerase, using forward primer Cre1 (5'-GGGAACGCACTTTCACAG-3'), reverse primer Cre2 (5'-CATCTACAAGCTGTACAGCTTC-3') and the following PCR parameters: denaturation at 94 ºC for 30 seconds, primer annealing at 55 ºC for 30 seconds, polymerization at 72 ºC for 60 seconds, for 35 cycles. The PCR product was

verified on a 2% agarose gel and purified with the QIAquick PCR Purification Kit. The probe was verified by cutting with BsmA I. Genomic DNA from 6-9x10^6 cells was then digested with Kpn I or Mfe I and analyzed with southern blotting as previously described.

3.2.3.7 Investigating Cellular Localization of AID.

eGFP-tagged AID DT40 cells were pre-treated with 10-25 ng/mL leptomycin B (LMB) for 2.5-5 hours prior to fluorescence microscopy. 15 μl of (in PBS buffer) cells were incubated with or without 0.5 μl/ml DRAQ5*TM* before mounting them onto poly-L-lysine coated microscope slides. The cells were investigated with DeltaVision de-convolution microscope using the Softworx 3.5 software and applying the following parameters: magnification=60-100x, eGFP-λ_{EX}=490 nm, eGFP-λ_{EM}=528 nm, DRAQ5^{TM}-λ_{EX}=555 nm, DRAQ5^{TM}-λ_{EM}=617 nm.

3.2.4 Methods related to identification of AID-interacting proteins.

3.2.4.1 Cross-linking antibodies to CNBr-activated Sepharose 4B.

The cross-linking of antibodies was carried out using CNBr-activated Sepharose 4B (GE Healthcare, UK) according to the manufacturer instructions using anti-Myc 9E10 at 10 mg antibody per 1 ml sepharose gel.

3.2.4.2 Cell fractionation in DT40.

DT40 cells were harvested at 1,200 rpm for 10 min, the supernatant was discarded and cell pellet washed with 50 ml cold PBS and repeated once. The whole procedure was carried out in the cold room (4 °C), except where indicated.

For cytoplasmic lysis, 5 times packed cell volume (1 ul PCV = 10^6 cells) equivalent of Isotonic Lysis Buffer [see 3.1.15, also in (Buerstedde and Takeda, 2006)] was added to the cell pellet, resuspended gently and then incubated for 12 min on ice to allow the cells to swell. To the swollen cells, 10% Triton X-100

(made into DT40 lysis buffer) was added to a final concentration of 0.3%, mixed thoroughly but gently and incubated for 3 min. The suspension was then centrifuged for 5 min at 1,500 rpm at 4 °C and the supernatant (cytoplasmic fraction) transferred to a fresh chilled tube. Prior to IP, the sample was centrifuged in ultracentrifuge with either a Ti 70.1 rotor for 2 min at 22,000 rpm at 4 °C or with a Ti 45 rotor for 2 min at 20,000 rpm at 4 °C. The supernatant was transferred to a pre-chilled tube, 1 x volume of cold DT40 Lysis Buffer was added to the supernatant in order to compensate the salt concentration. A sample from the cytoplasmic fraction was taken (50-80 ul) before adding affinity beads. The nuclear pellet was resuspended gently with a pipette in 5 x PCV equivalent of Isotonic Lysis Buffer and centrifuged at 3,000 rpm for 3 min. The supernatant was removed.

For nuclear lysis, nuclear pellets were resuspended in 2 x PCV DT40 Lysis Buffer+T (see 3.1.6) and dounce homogenized with 30-40 strokes in the cold room with a pre-chilled homogenizer. The samples were incubated with gentle agitation for 30 min at 4 °C and then centrifuged with a Ti 70.1 rotor at 22,000 rpm for 30 min at 4 °C or with a Ti 45 rotor for 30 min at 20,000 rpm at 4 °C (the rotor choice depended on the volume of the sample). The supernatants (nuclear fraction) was subjected to IP by adding affinity beads. Prior to this, a sample (50-80 ul) was taken.

The pellets were dounce homogenized in 2 x PCV DT40 Lysis Buffer+TB or DT40 Lysis Buffer+TD (if RNA was to be kept intact) or DT40 Lysis Buffer+TR (if RNA was to be degraded) per ml, until the pellets gave much less resistance. The samples were incubated at RT for 30 min until the disappearance of DNA fibers and centrifuged with either a Ti 70.1 rotor for 30 min at 22,000 rpm at 4 °C or with a Ti 45 rotor for 30 min at 20,000 rpm at 4 °C. The supernatants (chromatin fraction) were subjected to IP by adding affinity beads. Prior to this, a sample (50-80 ul) was taken.

3.2.4.3 Protein immunoprecipitation.

Anti-FLAG M2 affinity gel (Sigma, UK) was prepared as follows: 50% slurry affinity gel was washed with 10 x gel volumes of cold PBS at 4 ℃, centrifuged

for 3 min at 300 x g 3 times. The resin was resuspended in lysis buffer to make 50% slurry and added as follows for Cytoplasm - 3 μl of 50% slurry beads per $5x10^7$ DT40 endogenously tagged AID; Nucleoplasm - 1 μl of 50% slurry beads per $5x10^7$ DT40 endogenously tagged AID cells; Chromatin - 1 μl of 50% slurry beads per $5x10^7$ DT40 endogenously tagged AID. Samples were incubated for 3-4 hours at 4 ℃ on a rotator, centrifuged for 3 min at 300 x g at 4 ℃ and supernatants collected for analysis. The beads were washed 5 times with 25 x bead volume of DT40 Lysis Buffer (supplemented with 0.5-1 ug/ml 1xFLAG peptide) at 4 ℃ for 10 min (NaCl salt concentration wash increased up to 500 mM during IP optimisation experiments for increasing stringency). Between the washes, the samples were centrifuged for 3 min at 300 x g at 4 ℃ to pellet the resin. During the last wash, the supernatant with beads was transferred to a new tube, to remove non-specific proteins sticking to the tube walls. The beads were washed once more, this time in DT40 lysis buffer lacking Triton X-100. The protein was eluted twice by adding 4 x bead volume of elution buffer (DT40 Lysis Buffer + 500 μg/ml 3xFLAG peptide) and incubated for 30 min at 4 ℃ on a rotator. The samples were then centrifuged for 2 min at 2,000 rpm at 4 ℃ to pellet the resin. The supernatants were purified through a BioRad Bio-Spin Micro Chromatography Columns first by gravity flow and then by spinning 15 sec at 1,000 rpm at 4 ℃.

For large-scale experiments, tagged AID was eluted from the beads in BioRad Bio-Spin Micro Chromatography Columns, incubating it three times with 4 x bead volume of elution buffer for 30 min at 4 ℃ and leaving the last elution overnight. The eluted proteins were concentrated with StrataClean Resins, vortexing the mixtures for 1 min, and spinning them at 2,000 rpm for 1 min, followed by removal of the supernatants. The samples were stored at -20 ℃.

The sequences of peptides used in this study, were as follows:
3xFLAG peptide - MDYKDHDGDYKDHDIDYKDDDDK
1xFLAG peptide - DYDDDDK
3xMyc peptide - MEQKLISEEDLNEMEQKLISEEDLNEMEQKLISEEDLNE
1xMyc peptide – AEEQKLISEEDLL

3.2.4.4 Silver Staining.

The samples were run on a 4-12% NuPage gel, washed for 1 min in milliQ water. The gel was then subjected to silver staining using the SilverQuest Silver Staining Kit (Invitrogen Ltd, UK) according to manufacturer's instructions. The gel was washed with 100 ml water for 10 min and then dried by vacuum drying at 65 ºC for 1 hour.

3.2.4.5 SYPRO Ruby Protein Gel Staining.

The samples were run on a 4-12% NuPage gel, and washed for 1 min in milliQ water. Thereafter, the gel was subjected to silver staining using the SYPRO Ruby Protein Gel Stain (Molecular Probes, UK) according to manufacturer instructions. Before imaging, the gel was rinsed in ultrapure water twice for at least 5 min. The imaging was carried out using the Typhoon scanner or FujiFilm scanner.

3.2.4.6 Mass spectrometry.

Eluates from protein immunoprecipitation were subjected to gel electrophoresis by using 4-12% NuPage gels (Invitrogen, UK), followed by SYPRO Ruby staining and visualization by Dark Reader under UV. Individual protein bands were excised from the gel, and placed into separate 96-well plate wells in 100 μl of DEPC-treated H_2O. The samples were stored at 4 ºC until further processing. Peptides for analysis were generated by *in situ* tryptic digestion of protein/gel bands. Mass spectrometry analysis of peptides was performed on a SYNAPT HDMS mass spectrometer (Waters, USA) and the data searched against a concatenated, non-redundant protein database (UniProt 13.6), using the Mascot search engine (Matrix Science, USA). Mass spectrometry was carried out by Clare Hall and LRI mass spectrometry services.

3.2.5 Methods related to fission yeast experiments.

3.2.5.1 Fission yeast strains and media.

The genotypes of the *Schizosaccharomyces pombe* (*S. pombe*) strains used were JCF108 (h- *ade6*-M210 *his3*-D1 *leu1*-32 *ura4*-D18), JCF109 (h+ wt *ade6*-M216 *his 3*-D1 *leu1*-32 *ura4*-D18), F1693 (h- *rec12*-171::*ura4*+ *ura4*-D18 *leu1*-32 *arg3*-D4 *ade6*+), FO1694 (h+ *rec12*-171::*ura4*+ *ura4*-D18 *leu1*-32 his3-D1 *ade6*-M216), FO1695 (h- *rec12*-167 *leu1*-32 *his3*-D1 *arg3*-D4 *ade6*-M216) and FO1696 (h+ *rec12*-169::HA6HIS-KanMX *ura4*-D18 *leu1*-32 ade6+). *S. pombe* strains were stored in 2 ml of YES medium with 15% of glycerol at -80 ºC in cryotubes. For growth of fission yeast strains, a small amount of frozen glycerol stock was scraped with a pipette tip and then transferred to a YES plate. Strains were grown vegetatively in liquid/solid yeast extract with supplements (YE5S) or liquid/solid Edinburgh minimal medium (EMM) with supplements at 30 ºC as described previously (Moreno et al., 1991). Diploids were distinguished from haploid cells by adding 2.5 mg/liter phloxin B (Sigma) to the YE5S media or EMM+supplements. In the presence of phloxin B, diploid colonies stain dark pink whereas haploid colonies stain light pink.

3.2.5.2 Fission yeast mating.

Strains were crossed on EMM low nitrogen EMM low nitrogen supplemented with 225 mg/ml arginine, uracil and histidine but lacking leucine +/- thiamine (THIA) at 25 ºC. Strains from the opposite mating types (h+ and h-) were suspended in 50 µl of water, mixed and streaked onto the plates. The plates were incubated 2-3 days and analysed for the formation of asci by light microscopy.

3.2.5.3 Constructs for fission yeast experiments.

Exogenous expression of the human AID in fission yeast was achieved by using the pREP1 (nmt1) and pREP41 (nmt41) expression vectors described previously (Maundrell, 1990, 1993). These vectors contain the *nmt* (no message in thiamine) promoter, allowing efficient repression of gene expression by thiamine. A final

concentration of 15 μM thiamine was used to repress the *nmt* promoter. pREP41 has a weaker promoter compared to pREP1 and therefore less protein is produced from this vector.

AID, its catalytic mutant, and AID-eGFP fusion protein expression vectors were generated by cutting either pREP1 or pREP41 vector with Nde I and Bam HI restriction enzymes, run on an agarose gel and then purified from the gel. The same restriction sites were included in the oligonucleotides for amplifying the AID sequence. The inserts were cut with Nde I and Bam HI and ligated with the purified and restricted backbone. Apobec2 and Apobec3C expressing vectors were generated as follows: the Apobec2 and 3C sequences were PCR amplified, using primers with Nde I and Xba I sites. The PCR product was restricted with Nde I and Xba I sites and purified with PCR purification kit. His tag was obtained separately from two oligos with 5' Xba I and 3' Bam HI sites, which were hybridized, kinated and ligated with purified and restricted pRep1 backbone and with either purified Apobec2 or Apobec3C through a three fragment ligation process. The inserts of the expression vectors were confirmed by sequencing.

3.2.5.4 Lithium acetate transformation of fission yeast.

The protocol for lithium acetate transformation of fission yeast was used as described previously (Bahler et al., 1993). Cells were grown in 15 ml of YE4S liquid medium at 30 ℃ shaking until OD_{600} 0.2 – 0.5. Cells were pelleted at 3,000 rpm 5 min and supernatant was aspirated. The pellet was suspended in 1 ml water and transferred to a 1.5 ml tube. Cells were spun at 13,000 rpm 30 seconds, supernatant was removed and pellet was washed with 1 ml lithium acetate (LiAc)/TE. Cells were pelleted and resuspended in 1 ml LiAc/TE. 100 μl of cells were mixed with 50 μg boiled sonicated single stranded salmon sperm DNA. 0.5 μg of expression vector was added to the mixture and incubated 10 min RT. 280 μl 50% PEG LiAc/TE was added and incubated 45 min at 30 ℃. Thereafter, 43 μl DMSO was added and cells were heat shocked 5 min at 42 ℃. Cells were mixed with 500 μl, pelleted at 13,000 rpm 45 seconds and washed once with 1 ml water. Cells were resuspended in 50 μl of water and plated on EMM + supplements –Leu+THIA. Plates were incubated at 30℃.

3.2.5.5 Tetrad analysis.

Tetrad analysis was carried out as described (Moreno et al., 1991), with minor modifications. Asci from 3-day old matings were suspended in 0.5 ml of dH$_2$O. A drop of water with asci was run down the side of YE5S plates and left to dry at RT. Thereafter plates were incubated at 37 ºC for 4-5 hours until the asci walls were nearly disrupted, with the lid of the plate partially open. Spores from either tetrads, triads or dyads were individually picked and segregated with a micromanipulator. YE5S plates were incubated at 30 ºC until colonies become visible. These were counted to estimate the viabilities of spores from tetrads, triads and dyads.

Tetrad, triad and dyad frequencies were calculated by counting ethanol fixed asci under fluorescence microscopy. Mis-segregation was assessed in mature asci by the presence of more than 4 spores, less than two spores or asci in which there were clearly detectable aberrances in the DNA content according to DAPI staining and in the relative sizes of the spores.

3.2.5.6 Random spore analysis.

Approximately 1 x 10^5 cells from 3-day old matings were suspended in 250 μl of water containing 1500 units of β-glucuronidase (Sigma, UK) and were incubated overnight at 30 ºC to digest vegetative cells. Spores were pelleted for 20 seconds at 13,000 rpm, the supernatant aspirated and cells resuspended in 1 ml of dH$_2$O. Spores were counted with a haemocytometer, 1000 spores were grown on YE5S or YE5S with 15 μM of thiamine plates for 2-3 days at 30 ºC and corresponding total number of colonies was counted.

3.2.5.7 Recombination assay.

A loopfull of 3-day old yeast matings was treated with β-glucuronidase. Spores were counted with haemocytometer and plated (50,000 spores per plate) on selective media EMM-Leu-Ade-Ura+Phloxin B with or without thiamine and non-selective media EMM-Leu+Ade+Ura+Phloxin B with thiamine. Plates were incubated at 30 ºC for 5 days. Light pink colonies representing haploid cells were counted, dark pink colonies representing diploid cells were excluded from the

counting.

3.2.5.8 Analysis of GFP-AID expression during *nmt* promoter activation or repression.

To analyse the activation of *nmt* promoter, cells were grown vegetatively in 4 ml EMM-Leu+Thiamine at 30 ºC to an OD of approximately 1.0, centrifuged at 3000 rpm for 5 min, and the cell pellet washed with 1 ml of dH_2O. Cells were centrifuged again and strains were mixed to start the cross. The cross was carried out on EMM low nitrogen plates in the absence of thiamine at 25 ºC. Analysis was carried out by flow cytometry as described below.

To analyse the repression of nmt promoter by thiamine, cells were first vegetatively grown in the absence of thiamine and strains mixed together in water containing 15 μM thiamine for promoting efficient *nmt* repression. Thereafter, the cross was performed on EMM low nitrogen plates containing 15 μM thiamine. Every hour, a loopful of cross was fixed in 1 ml of cold 70% ethanol, vortexed and stored at 4 ºC. Cells were spun at 13,000 rpm for 30 seconds, the cell pellet washed in 1 ml 50 mM sodium citrate buffer pH 7.0, then harvested in 300 μl 50 mM sodium citrate buffer pH 7.0. Samples were analysed using a Beckton-Dickinson FACScan counting a total of 50,000 cells per sample.

3.2.5.9 Nocodazole induced block during meiosis.

Cells were grown vegetatively and crossed in the presence of thiamine for 24 hours in order to promote the fusion of the cells from opposite mating types. Thereafter cells were treated with nocodazole (final concentration 300 μM) for 24 hours to finish DNA replication of the fused cells and to block other cells from entering meiosis. Cells were washed extensively to remove thiamine, in the presence of nocodazole, to induce AID expression for 24 hours. GFP-AID induction was confirmed by flow cytometry. Nocodazole was removed to induce spore formation, while inhibiting AID expression with thiamine. Spores were freed from asci by β-glucuronidase treatment and plated on EMM media with thiamine lacking adenine, uracil and leucine to identify recombinants or EMM with thiamine lacking leucine to identify the overall numbers of live spores.

3.2.6 Methods related to nematode experiments.

3.2.6.1 Worm strains and culture conditions.

C. elegans strains were cultured as described previously (Brenner, 1974). Animals were cultured on MYOB plates seeded with *E. coli* and incubated at 15 ℃ or 22 ℃. The propagation of the animals was carried out by transferring approximately 1 cm^2 of agarose with a clean and sterilized scalpel onto freshly seeded MYOB plates. The strains wild-type Bristol N2 and *spo-11* (*ok79*) (Dernburg et al., 1998) were kindly provided by the *Caenorhabditis* Genetics Centre (University of Minnesota, St Paul, MN). The *ung-1* strain was kindly provided by Dr. Hilde Nilsen (University of Oslo) (Dengg et al., 2006).

3.2.6.2 Generation of human AID expressing worm strain.

Human AID sequence was cloned into the Gateway System (Invitrogen Ltd, UK) entry vector, p221, followed by cloning into pID3 vector generating an *AID::GFP* fusion protein under the control of the *pie*-1 promoter (Tenenhaus et al., 1998). The *AID::GFP* transgenic worm line was generated by microparticle bombardment (described below) of *unc-119* (*ed3*) animals (Praitis et al., 2001) with pID3_AID::GFP. *Spo-11 AID::GFP* worms were generated by crossing *unc-119 AID::GFP* worms with a GFP balanced *spo-11* strain. The genotype of the worms was verified by single worm PCR, using primers for AID and *spo-11* as described below.

3.2.6.3 Bleaching worms for bombardment.

Worms were grown on 10 plates until high density, washed off the plates with 10 ml M9 buffer, allowed to pellet, supernatant removed and washed again with 10 ml M9. 15 ml of bleach (1.5 ml bleach + 1 ml 5M NaOH + 12.5 ml dH$_2$O) was added and shaken vigorously. After 2 min, worms were passed through a 23G needle, and shaken and vortexed for 6-8 minutes, until most bodies were broken and eggs released. Eggs were washed 4-5 x in 50 ml dH$_2$O, spinning them down

at 1,500 rpm for 3 minutes between each wash. Eggs were then put into 10 ml S-basal medium and rotated overnight at RT to hatch the eggs and obtain L1 larva.

3.2.6.4 Preparing the gold beads for bombardment.

30 mg of 1 micron gold beads were weighed into a siliconized eppendorf tube. 1 ml of 70% ethanol was added and vortexed for 5-10 minutes. Particles were settled for 15 minutes and spun briefly (3-4 seconds). 1 ml of dH_2O was added to the particles, vortexed for 1 minute, allowed to settle for 1 minute and then spun briefly (3-4 seconds) in microfuge, this was repeated 3 times. Supernatant was then removed with vacuum pump, 250 μl of 50% glycerol was added, and stored at RT.

3.2.6.5 Loading DNA onto the gold beads.

Prepared beads were vortexed in 50% glycerol for 20 min. During vortexing, the following compounds per sample were added dropwise to the beads: 0.79 μg of plasmid, 10 μl 2.5 M $CaCl_2$, 4 μl 0.1 M freshly made spermidine. Between the addition of each compound, beads were vortexed for 1 minute, and finally vortexed a further 3 minutes. Beads were left to settle for 1 min and centrifuged briefly (3-4 seconds) in microfuge. Supernatant was removed using a pipetman to avoid disturbing the pellet of beads. Pellets were resuspended in 30 μl of 70% ethanol per bombardment and the ring of beads on the tube wall was scraped off with the pipet tip while adding the ethanol. Beads were centrifuged briefly, supernatant was removed and particles were resuspended in 10 μl of 100% ethanol per bombardment. Particles were vortexed at least 3 minutes and thereafter resuspending with vortexing and pipeting was continued until most of the large clumps of beads were broken up.

3.2.6.6 Putting beads on macrocarriers.

Macrocarriers were placed into the vacuum holders using a pen. Approximately 10 μl of the gold beads suspended in ethanol were spread over the centre of each

macrocarrier, avoiding the placing of beads outside the hole in the macrocarrier holders. Macrocarriers were allowed to dry during the preparation of worms.

3.2.6.7 Preparing plates with worms for bombardment.

Synchronized *unc-119 (ed3)* mutant L1 larvae were poured into 100 ml of S-basal + Nystatin + Streptomycin + 2 g of HB101. Worms were grown at 20 ºC with shaking at 250 rpm. After approximately 4 days, the culture produced a large number of L4/young adult worms, which are the most effective for bombarding. Worms were left to settle, washed with 40 ml of M9, and again left to settle for 15-20 minutes. Worms were resuspended in 5 ml of M9 and counted (10 x 2 µl to estimate the average worm number per µl). The animals were resuspended in M9 to 10,000 worms in 75 µl. Just before each bombardment, 75 µl of worms were added to the center of a seeded 10 cm plate for each bombardment, that the target area of the plate was covered with a monolayer of worms.

3.2.6.8 Worm bombardment and post bombardment care of the worms.

Worm microparticle bombardment was carried out using a Biolistic PDS-1000/He (BioRad, UK) with 1,350 psi rupture discs. Target plates with worms were placed on the second shelf from the bottom and a vacuum of 25 inches Hg was used during bombardment.

After bombardment, worms were left to recover for 1-2 hours, washed off from the plates with S-basal, allowed to settle, resuspended in 400 µl of S-basal per bombardment, and seeded onto 4 freshly seeded 10 cm plates to isolate the maximum number of independent lines. Plates were incubated at RT for 2 weeks and then screened for transformants. The transformant are rescued for *unc-119* allowing the worms to move on the starved plates. From each plate with transformants, 6 individual animals were picked and placed to a separate well of a freshly seeded 24 well plate. These were grown and the progeny was monitored for *unc-119* animals. Stable lines that did not produce *unc-119* animals were analysed further.

3.2.6.9 Single worm PCR.

Five microlitres of lysis buffer [50 mM KCl, 10 mM Tris (pH 8.3), 2.5 mM MgCl$_2$, 0.45% IGEPAL CA 630, 0.45% Tween-20, 0.01% Gelatin] with proteinase K (1 mg/ml) was placed into 0.2 ml DNase and RNase free PCR tubes or a 96-well PCR plate. Single worms were washed in a drop of PBS and then transferred into lysis buffer and frozen at - 80 ℃ or on dry ice for 20 min. Tubes were then incubated at 65 ℃ degrees for 90 minutes to lyse the worms. Proteinase K was inactivated by heating the tubes to 95 ℃ for 15 min. Thereafter, a nested PCR reaction was performed. For the first PCR reaction, 1 μl of worm lysate was added to a 12.5 μl PCR reaction that included the primers at 0.5 μM, 1x buffer, 1 mM MgCl$_2$, 0.2 mM dNTP's and 1.25 U of Platinum Taq DNA polymerase. The PCR conditions were as follows: denaturation at 94 ℃ for 30 seconds, primer annealing at 55 ℃ for 30 seconds, polymerization at 72 ℃ for 90 seconds, for 30 cycles. The second PCR reaction was performed using 1 μl of the first PCR reaction in a 12.5 μl PCR reaction mixture as indicated for the first reaction. The conditions for the second PCR were as follows: denaturation at 94 ℃ for 30 seconds, primer annealing at 55 ℃ for 30 seconds, polymerization at 72 ℃ for 90 seconds, for 25 cycles. AID-GFP fusion construct was detected by nested PCR using external primer sequences AID-GFP1F (5'-TAGACCCTG-GCCGCTGCTACC-3') and AID-GFP1R (5'-CAAAAGGATGCGCCG-AAGCTGTCTGGAG-3') and nested primers AID-GFP2F (5'-GAGGCAAG-AAGACACTCTGG-3') and AID-GFP2R (5'-GTGACATTCCTGGAAGTTGC-3'). *Spo-11* deletion was detected with nested PCR using external primers Spo11-1F (5'-CGTGTTTCCCAAGATGCTC-3') and Spo11-1R (5'-CGGAATGCG-TGCAAGTG-3') and nested primers Spo11-2F (5'-CCGAACAGCA-TATTGAAGAGG-3') and Spo11-2R (5'-GCGCATATAAAACACGGAGAC-3'), as previously described (Dernburg et al., 1998).

3.2.6.10 Slide preparation for immunofluorescence of *C. elegans*.

Slides were washed in 70% ethanol, before smearing them with 1 mg/ml poly-L-lysine for 5 min. Slides were left to dry at 37 ℃ and then given another coat of poly-L-lysine.

3.2.6.11 Visualisation of GFP in *C. elegans*.

Molten 4% agarose was placed on prewarmed (55 °C) slides and pressed between two slides to a final thickness of approximately 1 mm. The agarose was left to solidify before 5 worms were placed on the agar surface in 10 µl of PBS. Slides were left to dry for 5 min before covering with a coverslip. Slides were analysed using a Deltavision microscope (Applied Precision, UK).

3.2.6.12 Cytological preparation and immunostaining of *C. elegans*.

Gravid hermaphrodites were transferred to 30 µl of PBS on poly-L-lysine-coated slides. Thereafter, worms were transferred to 50 µl of 10 mM levamisole. Germlines were extruded by removing the head and tail with a fine needle (27-gauge). Levamisole was replaced with 1% paraformaldehyde in PBS for 10 min. After fixation, the germlines were permeabilised for 5 min in TBSBT (Tris-buffered saline, 0.5% bovine serum albumin, 0.1% Triton X-100) and washed with TBSBT at least three times for 5 min each time, followed by blocking in TBSBT for 30 min. Primary antibodies were diluted in TBSB [1:50 for both 4.18.1 and 4.26.1 (Aoufouchi et al., 2008) anti-AID, 1:200 for anti-UNG (ab13668, Abcam, Cambridge, UK), 1:200 for anti-RAD-51 antibody (Martin et al., 2005)] and incubated overnight at 4°C in a humid chamber. Germlines were subsequently washed at least three times for 5 min each time with TBSB (Tris-buffered saline, 0.5% bovine serum albumin) before incubation with secondary antibodies conjugated with Cy3 (anti-rabbit, 1:10,000, Sigma, UK; anti-mouse, 1:10,000; Sigma, UK) for 1 to 2 h at RT. Finally, the germlines were washed at least three times with TBSB for 5 min before adding ProLong Gold Antifade with DAPI (Molecular Probes, UK) and mounted with coverslips.

3.2.6.13 Fluorescence microscopy.

Delta-vision microscopy was used to examine germlines with x40 or x63, 1.4 NA Planapochromat lens on an Olympus inverted microscope (IX71), and images captured using SoftWorx computer software (Applied Precision, UK). Three-dimensional data sets were computationally deconvolved if needed, and regions

of interest then projected into one dimension. Merged or single color images were recorded using GIMP software.

3.2.6.14 Worm viability analysis.

Single L4 stage larvae were picked into individual seeded 24 well plates. After an initial 12 h egg-laying period, worms were transferred to fresh plates and once again transferred 12 h and 24 h later, so that progeny could be scored over one egg-laying period per parent. Plates were examined after 48 h for the presence of viable adult progeny. Unhatched eggs and viable progeny were counted to estimate the viability of the offspring.

3.2.6.15 Hermaphrodite / male identification.

The percent of male progeny was determined by sex-specific counting, which included the rapid movement of the animal, the presence of a hook-shaped tail, and slim torso. Statistical analysis of worm viability was done using chi-square analysis, with the following addition: in the wild-type worm we arbitrarily set the count of dead eggs to 2, even though we did not observe any. This alteration was to facilitate statistical analysis (being even more conservative than the observed data).

3.2.6.16 Preparation of PEG/DMSO competent cells.

HT115(DE3), an RNase III-deficient *E. coli* stain suitable for RNAi experiments, was streaked on an LB plate and incubated overnight at 37 ℃. A single colony was inoculated in 5 ml LB and incubated overnight at 37 ℃. This starter culture was used to inoculate 500 ml LB culture, and grown to OD600 = 0.5-0.6 at 37 ℃. The cells were pelleted at 2000 rpm for 5 min at 4 ℃ and gently suspended in 25 ml cold TSB. The cells were incubated on ice for 10 min, aliquoted, frozen and stored at -80 ℃.

3.2.6.17 Chemical transformation of bacteria for RNAi experiments.

Ung-1 sequence was Gateway cloned into pL4440, a modified version of

Bluescript vector carrying tandem T7 promoters on each side of the multi-cloning site, driving transcription of each DNA strand. The constructs were transferred (described above) into HT115(DE3) bacteria and plated onto LB plates containing ampicillin. DNA (pL4440 containg *ung-1* seqence or empty control plasmid) was mixed with 5xKCM and ddH$_2$O to a final volume of 100 µl at 1xKCM final concentration. PEG/DMSO competent cells were thawn on ice, added to the mixture in an equal volume, incubation on ice for 20 min, incubated at RT for 10 min, 1 ml of SOC media was added and incubated at 37 ºC for 1 h, and plated on LB plates with ampicillin and incubating at 37 ºC overnight.

3.2.6.18 RNA interference (RNAi) by feeding.

RNAi by feeding for *ung-1* (Y56A3A.29a) was performed for 24 h, as described previously (Boulton et al., 2002; Dengg et al., 2006; Fire et al., 1998). A colony with RNAi vector was grown for 6-18 hours in LB + 50 µg/ml ampicillin, seeded onto 6 well plates with NGM Classic media and incubated overnight at RT with 4 mM to allow RNAi production by IPTG. Then, L4 stage hermaphrodites were transferred onto separate plates, while minimizing the amount of the bacteria carried over from the MYOB plate. The plates were incubated 24-48 hours at RT and the offspring analysed for effects of RNAi. Empty vector (pL4440) was used as a control for RNAi experiments.

3.2.6.19 Irradiation of worms.

Irradiation of nematodes was performed in an [137]Cs Irradiator (IBL437, Cis Bio International) with low dose (15 Gy) irradiation. Animals were analysed for TUNEL staining 30 min post-irradiation and for RAD-51 staining 4 h post-irradiation.

3.2.6.20 TUNEL assay.

The TUNEL (terminal deoxynucleotidyl transferase-mediated dUTP nick end-labelling) assay was performed as described earlier (Parusel et al., 2006), with minor modifications. Gravid nematodes were washed in PBS on poly-L-lysine

covered slides and dissected in PBS, followed by fixation in 4% formaldehyde in PBS at RT for 20 min. Gonads were rinsed in PTX (PBS, 0.4% Triton X-100) three times, incubated in 100 mM sodium-citrate, 0.1% Triton X-100 at 65 ℃ for 20 min and then washed twice in PTX. Gonads were incubated for 30 min at RT in 0.1M Tris/HCl (pH 7.5) containing 3% BSA and 20% bovine serum, washed twice with PBS at RT and excess fluid was drained off. Thereafter, the gonads were incubated in TUNEL reaction mixture (Roche, USA) at 37 ℃ for 1.5 hours in the dark. The reaction was stopped by washing three times with PTX, before adding ProLong Gold Antifade with DAPI (Molecular Probes, UK) and mounting the samples with coverslips.

Chapter 4

RESULTS

4.1 Hormonal regulation of AID.

4.1.1 Introduction.

4.1.1.1 Development of B-cells.

The induction of AID expression during B-cell activation is essential for Ig diversification and plays an important role in humoral immune responses. These immune responses are initiated when antigen binding B-cells receive co-stimulatory signals, which can include cytokines, direct T-cell help, complement or pathogen antigens. B-cell activation can be mimiced in tissue culture conditions by using various cytokines and bacterial cell wall components. *In vivo*, B-cell activation by T helper cells occurs in secondary lymphoid tissues, where the B-cell together with the T-cell establishes a primary focus of clonal expansion (Abbas, 2007). After several days, some of the proliferating B-cells differentiate into antibody-synthesizing plasma cells. Other B-cells migrate into a primary lymphoid follicle where they continue to proliferate and ultimately form a germinal centre, composed mainly of proliferating B-cells and antigen-specific T-cells. The early events in the primary focus lead to the prompt secretion of specific antibody that serves as immediate protection to the infected individual. The germinal centre reaction, on the other hand, provides for a more effective later response, should the host become infected again. In germinal centers, B-cells undergo several important processes (McHeyzer-Williams et al., 2006), including somatic hypermutation of V regions, affinity maturation that selects for survival of B-cells with high affinity for the antigen, and isotype switching that allows B-cells to express a variety of effector functions in the form of antibodies, processes described earlier (Chapter 1).

4.1.1.2 Gender differences in immune responses and autoimmunity.

During an infection, the organism needs to trigger an immune response to eliminate the source of the infection, and at the same time, avoid harming its own tissues. Interestingly, there is a sexual dimorphism in the immune response in

humans, which was noted already several decades ago, indicating that females exhibit a quicker and stronger humoral immune response upon infection than males (Butterworth et al., 1967). Similarly, after immunization, female mice produce more antibody than male mice (Eidinger and Garrett, 1972; Weinstein et al., 1984). Although a component of this sexual dimorphism in immunity is likely due to genetic differences in mouse models, systemic sex hormone treatment or castration have been shown to switch the nature of immune responses to that of the opposite sex. This would suggest that the sex differences are due, in part, to differential hormone levels *in vivo* (Bebo et al., 2001; Bebo et al., 1999).

There are more than 70 autoimmune diseases affecting approximately 5 % of the population of developed countries, and in some of these pathologies, abnormalities in B-cells seem to play an important role (Davidson and Diamond, 2001). An autoimmune disease occurs when a specific adaptive immune response is mounted against self antigens (Ermann and Fathman, 2001). When such an event occurs, the immune effector mechanisms are unable to eliminate the antigen completely, so a sustained response occurs. This causes a chronic inflammatory injury to affected tissues. Cytotoxic T-cell responses and inappropriate activation of macrophages by Th1 cells can cause extensive tissue damage, whereas inappropriate T-cell help to self-reactive B-cells can initiate harmful autoantibody responses (Davidson and Diamond, 2001; Taneja and David, 2001).

The susceptibility to autoimmune diseases is controlled by both environmental and genetic factors (Ermann and Fathman, 2001; Goodnow, 2007). Firstly, several diseases are associated with the major histocompatibility complex (MHC) genotype. Secondly, in diseases such as Systemic Lupus Erythematosus (SLE), the disease susceptibility genes include genes that regulate the thresholds for tolerance and activation of B-cells [*e.g.* tumor suppressor Fas (CD95) and Fas ligand (Rieux-Laucat et al., 1995)]. Fas is a central protein for inducing apoptosis of autoreactive lymphocytes during negative selection in lymphocyte development, and elimination of autoreactive B-cells in the periphery (Rieux-Laucat et al., 2003; Rieux-Laucat et al., 1995). Thirdly, an important factor in disease susceptibility is the hormonal status of the patient (Christen and von Herrath, 2004; Janeway, 2001; Kamradt and Mitchison, 2001; Whitacre, 2001).

Female sex biases of over 80% have been reported in autoimmune diseases like Sjogren's syndrome, autoimmune thyroid disease (Hashimoto's thyroiditis and Graves' disease), SLE and scleroderma (Whitacre, 2001). Other relatively common diseases including myasthenia gravis, rheumatoid arthritis (RA) and multiple sclerosis (MS) have a sex distribution of 60–75% women (Ermann and Fathman, 2001; Whitacre, 2001). The increased prevalence of autoimmune disease in women, the sexual dimorphism of the immune response, and the modulatory effects of sex steroids on immune function *in vitro* suggest their involvement in mediating the sex differences in immune responses.

4.1.1.3 Systemic lupus erythematosus.

One of the better characterized autoimmune diseases, which shows a striking gender bias in respect of its occurrence (90% are female), is Systemic Lupus Erythematosus. SLE is a polygenic autoimmune disease that is characterized by a breakdown of B-cell tolerance, which results in development of autoantibodies (Cohen-Solal et al., 2006; Grimaldi et al., 2005). In addition to its strong prevalence in women, it has also a characteristic age onset after menarche and before menopause (Cervera et al., 1993). Outside the period of female reproductive activity, the onset of disease is uncommon and without sex preference (Font et al., 1991; Tucker et al., 1995). These observations suggest that endogenous sex hormones may play a role in the development of the disease, with estrogen acting to trigger disease susceptibility. Consistent with this, women with SLE tend to have low levels of plasma progesterone (Arnalich et al., 1992; Lahita et al., 1979) and abnormal patterns of estradiol metabolism, leading to increased estrogenic activity (Lahita et al., 1979). Exogenous estrogen therapies, hormone replacement therapy (HRT), and oral contraception (OCP), have also been associated with SLE disease flares (Sanchez-Guerrero et al., 1995). Murine models that develop a syndrome resembling human SLE (NZB/NZW mice) demonstrate that estrogen can modulate this disease (Roubinian et al., 1979). Female mice treated with estrogen (17β-estradiol) manifest an earlier onset of lupus and an earlier mortality (Roubinian et al., 1979). Conversely, when female mice are ovariectomized and treated with antiestrogenic drug such as tamoxifen, they exhibit a prolonged lifespan (Sthoeger et al., 2003).

Since in SLE, breakdown of B-cell tolerance and the production of autoreactive antibodies play an important pathological role, AID is a potential candidate involved in this process. Interestingly, over-expression of AID in autoimmune prone mice leads to a more severe SLE-like phenotype (Hsu et al., 2007), while breeding AID deficient mice with autoimmune prone MRL/lpr mice significantly reduces the on-set and extent of disease (Jiang et al., 2007), indicating that alterations in AID can change the severity of B-cell autoimmunity. Due to a gender bias in immune responses and autoimmune pathologies, it is therefore possible that AID is a target for hormonal regulation.

4.1.1.4 Estrogen and progesterone receptors.

Estrogen receptor (ER) and progesterone receptor (PR) belong to the nuclear steroid receptor family. They mediate the effects of sex hormones estrogen and progesterone by binding their ligands. Estrogen and progesterone hormones are synthesized from a cholesterol backbone and produced predominantly in the ovary. Besides having a central role in reproductive and developmental events, they are also involved in the modulation of the immune system (Li et al., 2004). ER and PR have distinct structural and functional domains: (a) an N-terminal region with a transactivation function, (b) a zinc finger region that mediates DNA binding, and (c) a C-terminal ligand-binding moiety (Sonoda et al., 2008). Two types of estrogen receptor, estrogen receptor α (ERα) and estrogen receptor β (ERβ) mediate the effects of estrogen (Bjornstrom and Sjoberg, 2005). These receptors are expressed not only in reproductive tissues but also in multiple other cell types, including cells of the immune system, such as B and T lymphocytes (McMurray et al., 1991; Peeva et al., 2005; Tornwall et al., 1999). However, there are differences in the structure and cellular distribution of ERα and β that suggest different biological roles for these two receptors (Gruber et al., 2002). In addition to intra-cellular receptors, there are also membrane-associated estrogen receptors that are involved in non-genomic rapid responses to hormones (Benten et al., 2002). Similarly to ER, PR has also 2 isoforms, PRA and PRB, which share several structural domains, but are distinct transcription factors that mediate their own response genes and physiologic effects (Leonhardt et al., 2003).

After estrogens bind to the ligand domain of the ER, the protein undergoes conformational changes that allow interactions with coactivator molecules and translocation into the nucleus (Tsai and O'Malley, 1994). In the classical pathway of estrogen action, the ER binds as a dimer to estrogen response elements (EREs) in the promoter regions of target genes, acting as an activating or repressing transcription factor (Beato and Klug, 2000; Tsai and O'Malley, 1994). However, it has been found that ER can regulate genes that lack a canonical ERE, suggesting additional pathways for estrogen action. For example, the ER may activate genes via protein tethering with c-Fos/c-Jun B (AP-1) and Sp1 (Kushner et al., 2000; Saville et al., 2000) or cross-talk with NFκB pathways (Ray et al., 1994).

The signaling pathway mediated by progesterone and its receptors resemble estrogen-mediated pathways. Upon hormone binding, the receptor dissociates from the heat shock protein complex, dimerizes, and binds to progesterone-responsive elements (PREs) within the regulatory regions of target genes. This leads to induction or silencing of target genes (Leonhardt et al., 2003). Similarly to ER, when bound to DNA, the PR dimer contacts components of the general transcription machinery, either directly (Ing et al., 1992; Kalkhoven et al., 1996) or indirectly via coactivators and corepressors (Onate et al., 1995; Voegel et al., 1996).

In addition to activated B-cells, AID expression has been reported in some hormone-responsive tissues outside the immune system (Morgan et al., 2004; Schreck et al., 2006), raising the possibility that AID expression might be regulated not only by classical B-cell transcription factors such as NFκB (Dedeoglu et al., 2004), Pax5 (Gonda et al., 2003) and E-box proteins (Sayegh et al., 2003), but also by hormones such as estrogen. We therefore set forth to test, whether AID expression could be affected by estrogen and progesterone.

4.1.2 Results

4.1.2.1 Differential effect of sex hormones on AID mRNA.

The regulation of AID during B-cell activation is mediated by various signaling pathways. To investigate the effects of estrogen and progesterone on AID induction in B-cells, murine splenic B-cells were isolated by negative isolation and stimulated with IL4 and LPS – known to induce AID (Dedeoglu et al., 2004). As hormones, including estrogen and progesterone, exert their biological effects predominantly through binding to their intracellular receptors and act as transcription factors, we examined AID mRNA levels upon B-cell stimulation and progesterone or estrogen treatment.

Addition of physiological amounts of progesterone during B-cell stimulation led to a 5-fold decrease in AID mRNA levels compared to the control DMSO treatment, as revealed by quantitative real time PCR (Figure 11A). In contrast to progesterone, physiological amounts of estrogen were able to enhance AID expression three fold in these cells (Figure 11A). To confirm that the estrogen induced stimulation on AID was analogous to other known estrogen response genes, we observed estrogen induction of GREB1, a known estrogen-responsive gene (Ghosh et al., 2000) (Figure 12). These results indicated that estrogen as well as progesterone can participate in the regulation of AID mRNA in activated B-cells in a stimulatory or inhibitory manner, respectively.

To investigate whether estrogen could up-regulate AID mRNA without IL4/LPS stimulation, isolated B-cells were treated with estrogen in the absence of LPS and IL4. Surprisingly, estrogen treatment led to 7 fold up-regulation of AID message in mouse B-cells (Figure 11B). This induction began to plateau at 4 h, with the earliest indication of an increase detectable after 1-2 h (Figure 11C). Since it was possible to enhance AID mRNA production with estrogen after co-treatment with IL4/LPS, it indicated a possible synergy between these two signaling pathways.
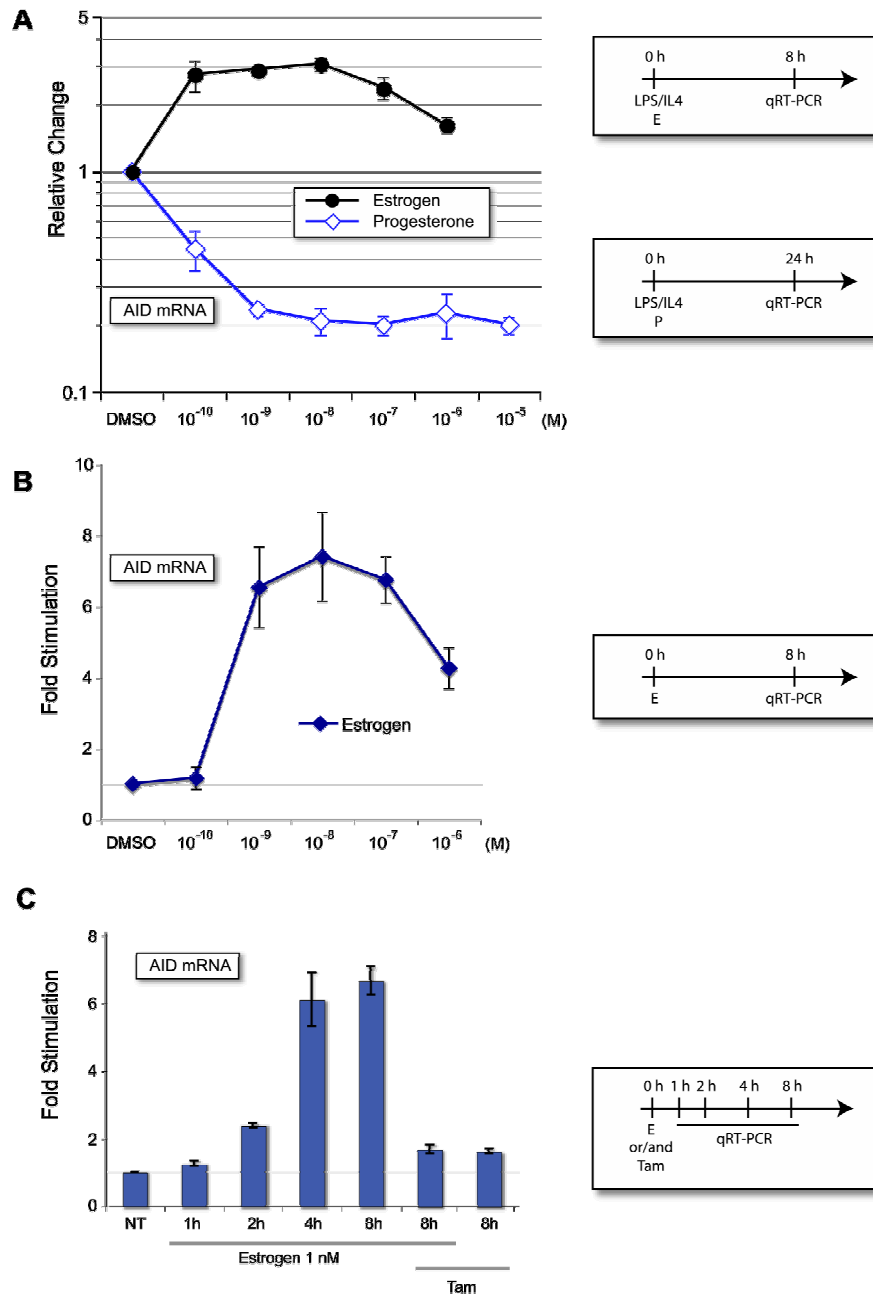
Figure 11. The effects of estrogen and progesterone on AID mRNA in splenic B-cells. (A) Changes in AID mRNA in response to estrogen and progesterone treatment in stimulated B-cells. Isolated mouse splenic B-cells were stimulated with LPS and IL4 and treated with different physiological concentrations of estrogen for 8 h or progesterone for 24 h. (B) AID mRNA in response to estrogen treatment in un-stimulated B-cells, after 8 h treatment with physiological concentrations of estrogen. (C) AID mRNA induction upon different treatments. Cells were stimulated with estrogen up to 8 h. Data is representative of three independent experiments and error bars indicate standard deviations from the average. DMSO is set to 1, and treatments are represented as relative change to DMSO. Time-lines of cell treatments are indicated next to the graphs. NT - not treated; Tam – tamoxifen (50 nM). Published in (Pauklin et al., 2009b).

Figure 12. The effects of estrogen on GREB1 mRNA in splenic B-cells. Unstimulated splenic B-cells were treated with 10 nM estrogen for 4 h or 24 h.
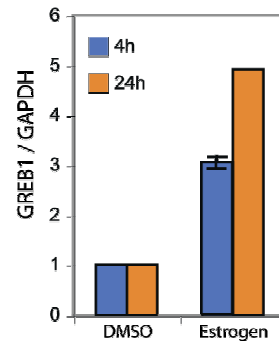
Tamoxifen, a molecule that has been used as an antagonist for estrogen in treating breast cancer, attenuated the effects of estrogen on AID induction (Figure 11C), but caused itself a modest increase in AID mRNA at low concentrations (Figure 13). Therefore, this structural analogue of estrogen is not a clear-cut antagonist of estrogen, but also a partial agonist in respect to AID induction.
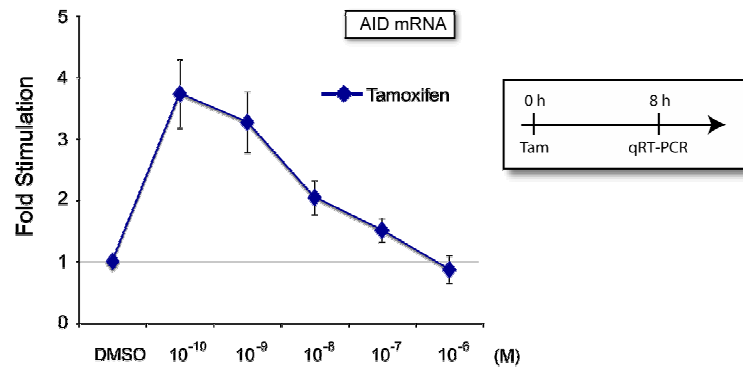
Figure 13. The effects of tamoxifen on AID mRNA in splenic B-cells. Relative changes in AID mRNA in unstimulated B-cells in response to 8 h treatment with various concentrations of tamoxifen. Published in (Pauklin et al., 2009b).

4.1.2.2 AID mRNA is regulated by hormones predominantly via transcription.

Since modest effects of hormones were already observed after 2 h (Figure 11C), it was necessary to clarify whether these changes are due to a direct effect on AID or mediated via a multistep pathway. To address this, splenic B-cells were pre-treated with translational inhibitor (cycloheximide - CHX) followed by

stimulation with estrogen. This did not alter the increase of AID mRNA production (Figure 14A), suggesting that the effect of estrogen was exhibited on AID's mRNA synthesis. qRT-PCR of cDNA is a read-out of steady state mRNA – therefore it was further investigated, whether the effects on AID mRNA by hormones were due to changes in transcription or mRNA splicing. Treatment of cells with transcription inhibitors α-amanitin (AMA) and actinomycin D (ACT) diminished the effect of estrogen, indicating that this alteration in AID's mRNA was not due to message stability (Figure 14A). Since estrogen is known to affect pre-mRNA to mRNA processing (Auboeuf et al., 2002), qRT-PCR primers were designed to span the complete transcription unit of AID. When the relative change of expression of the various exon and intron units of the pre-mRNA were compared to that of the mature mRNA, no major changes were observed (Appendix Figure 1B). However, the intron between exon 3 and 4 (7538-7689 bp - E) showed a marginal increase in response to estrogen over that of the mature RNA, indicating a potential region for estrogen induced splicing of AID RNA. As this increase was modest, it was not analysed further, although there is evidence that AID can be alternatively spliced (Wu et al., 2008).

Since the above data indicated that estrogen was likely to function on the level of transcription, we also determined, whether progesterone utilized an analogous pathway. In order to analyse the repressive effects of progesterone, we stimulated B-cells with IL-4 and LPS prior to hormone treatment. In contrast to estrogen, progesterone treatment of stimulated splenic B-cells also resulted in a decrease in AID mRNA production independently of translation (Figure 14B), mRNA stability (Figure 14B) or pre-mRNA splicing (Appendix Figure 1C).
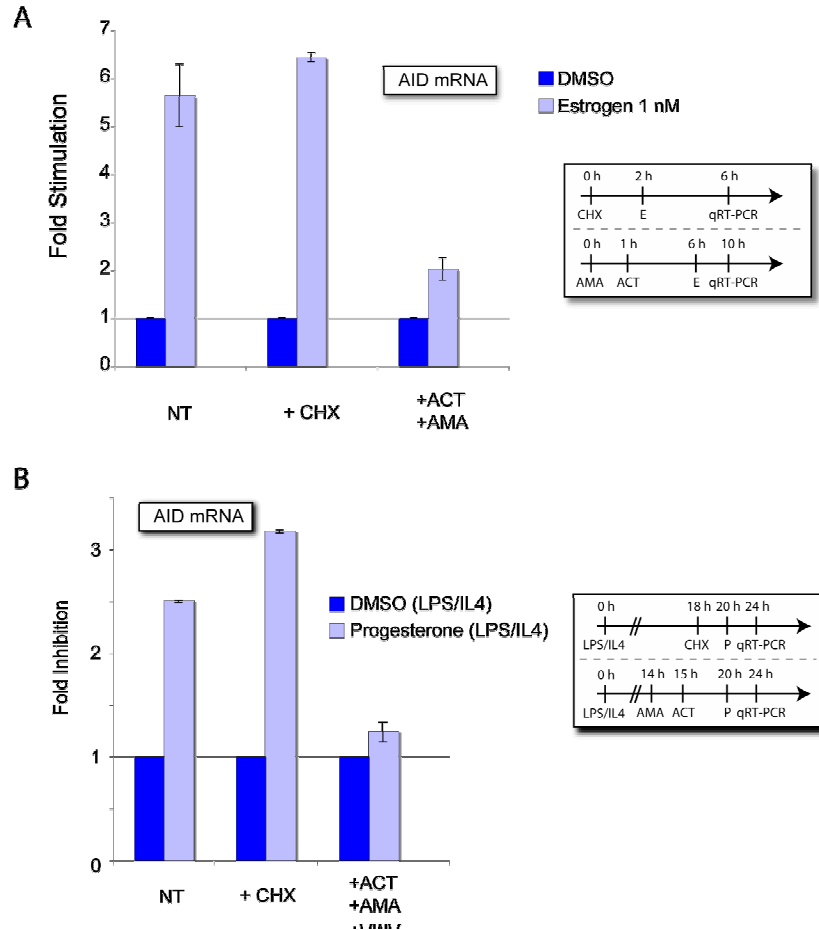
Figure 12. Estrogen and progesterone affect AID transcription directly. (A) Isolated mouse splenic B-cells were treated with CHX for 2 h or with ACT for 6 h and AMA for 5 h, followed by 1 nM estrogen treatment for 4 h, and AID mRNA was analyzed by qRT-PCR. Gene expression was normalized to the control treatments with DMSO. Published in (Pauklin et al., 2009b). (B) Isolated mouse splenic B-cells were first stimulated with LPS and IL4 and then treated with CHX for 2 h (18 h post stimulation) or with ACT for 6 h (14 h post stimulation) and AMA for 5 h (15 h post stimulation), and 200 nM progesterone (20 h post stimulation) for 4 h. Time-lines of cell treatments are indicated next to the graphs. NT - not treated. Published in (Pauklin and Petersen-Mahrt, 2009).

4.1.2.3 Identification of hormone response elements in the AID promoter regions.

The results suggested a direct regulation of AID by hormones. We therefore identified potential EREs (estrogen response elements) and PREs (progesterone response elements) in the promoter of human AID and other DNA daminases, using bioinformatic analysis (Figure 15). This indicated potential EREs and PREs in human AID promoter region and putative response elements in mouse

Apobec3 and human APOBEC3F and -3G.
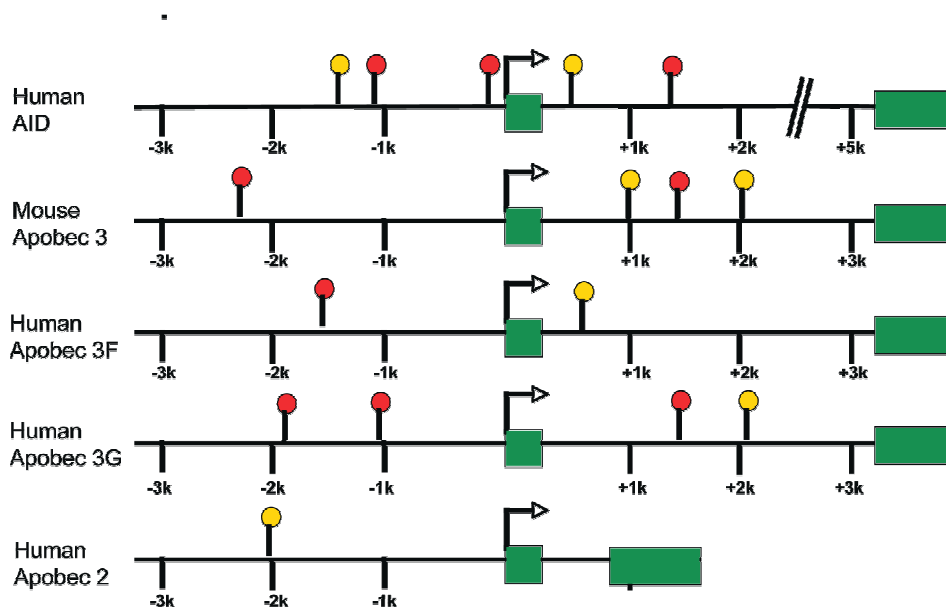
Figure 13. Schematic representation of the putative hormone response elements in DNA deaminase promoter regions. The promoter regions were analysed by bioinformatics using AliBaba2.1 transcription factor response element predicting program and Transfac 4.0 transcription factor binding site database, to identify potential estrogen response elements (marked with red) and progesterone response elements (marked with yellow) in the proximity of DNA deaminase promoter regions. Arrow depicts the transcription start site and green boxes depict exons. +/-k marks the distance in kilobasepairs downstream or upstream of transcription start site.

To examine which of these elements is functional, we dissected the AID promoter 1.5 kb upstream and 2 kb downstream of the ATG region and analysed these fragments for hormone responsiveness in a heterologous transcription assay. For this, promoter fragments were PCR amplified, inserted into a luciferase reporter construct containing a minimal promoter and transfected into SiHa cervical cancer cells (Figure 16A). Thereafter, cells were treated with the indicated hormones (or co-transfected with expression plasmids), and analyzed for luciferase activity. The results were normalised to control treatments with DMSO. Estrogen responsiveness could be seen on the upstream (Fragment C) region (Figure 16B). Since we did not detect any changes to estrogen at the most ATG proximal ERE (Fragment A) or Fragment E, these fragments were not studied further. However, these results did not rule out the possibility that the putative ERE near the promoter is functional in other cell lines, as there can be

heterogeneity in the responsiveness to estrogen in different cell lines (Cerillo et al., 1998) or a potential cooperation between the proximal and distal putative EREs on Fragment C. Additionally, depending on the stimulus, a cooperative action between transcription factors might exist *in vivo* (see below).

Aside from a putative ERE, Fragment C also harbored the two published NFκB binding sites (Dedeoglu et al., 2004). As already indicated by the qRT-PCR analysis (Figure 11A), estrogen and NFκB pathway (LPS/IL4) could act synergistically on AID mRNA production. The co-treatment of the SiHa cells with estrogen and TNFα (known to activate the NFκB pathway), resulted in a synergistic induction of the AID promoter, compared to TNFα or estrogen treatment alone. Interestingly, the effect was only detected at low estrogen concentrations, ranging from 0.1 nM to 10 nM (Figure 16C). This effect was abolished at higher estrogen concentrations, and even indicated a modest decrease, suggesting a complex interplay of these signals. To see whether NFκB and estrogen pathways can activate AID separately from each other, cells transfected with Fragment C construct were treated with TNFα and estrogen, while co-transfected with the dominant negative mutant of IκBα (S32A/S36A - IκBα-mt); a mutant that inhibits the release of NFκB from the cytoplasm into the nucleus after stimulation. Although TNFα mediated activaton of NFκB was inhibited by IkBα-mt as expected, estrogen was able to independently activate the transcription (Figure 16D).

To identify possible PREs, cells were stimulated with TNFα and progesterone followed by the analysis of luciferase signal. TNFα was used to activate NFκB, because we would not be able to examine the inhibitory effects of progesterone without prior stimulation of AID expression. When compared to DMSO treatment, a decrease in luciferase activity was detected upon progesterone treatment on the fragment containing the NFκB sites (Fragment C – Figure 16E). This indicates that the TNFα and hormone response elements in the AID promoter can either act in synergistically or inhibitory fashion, as for ERE and PRE, respectively.
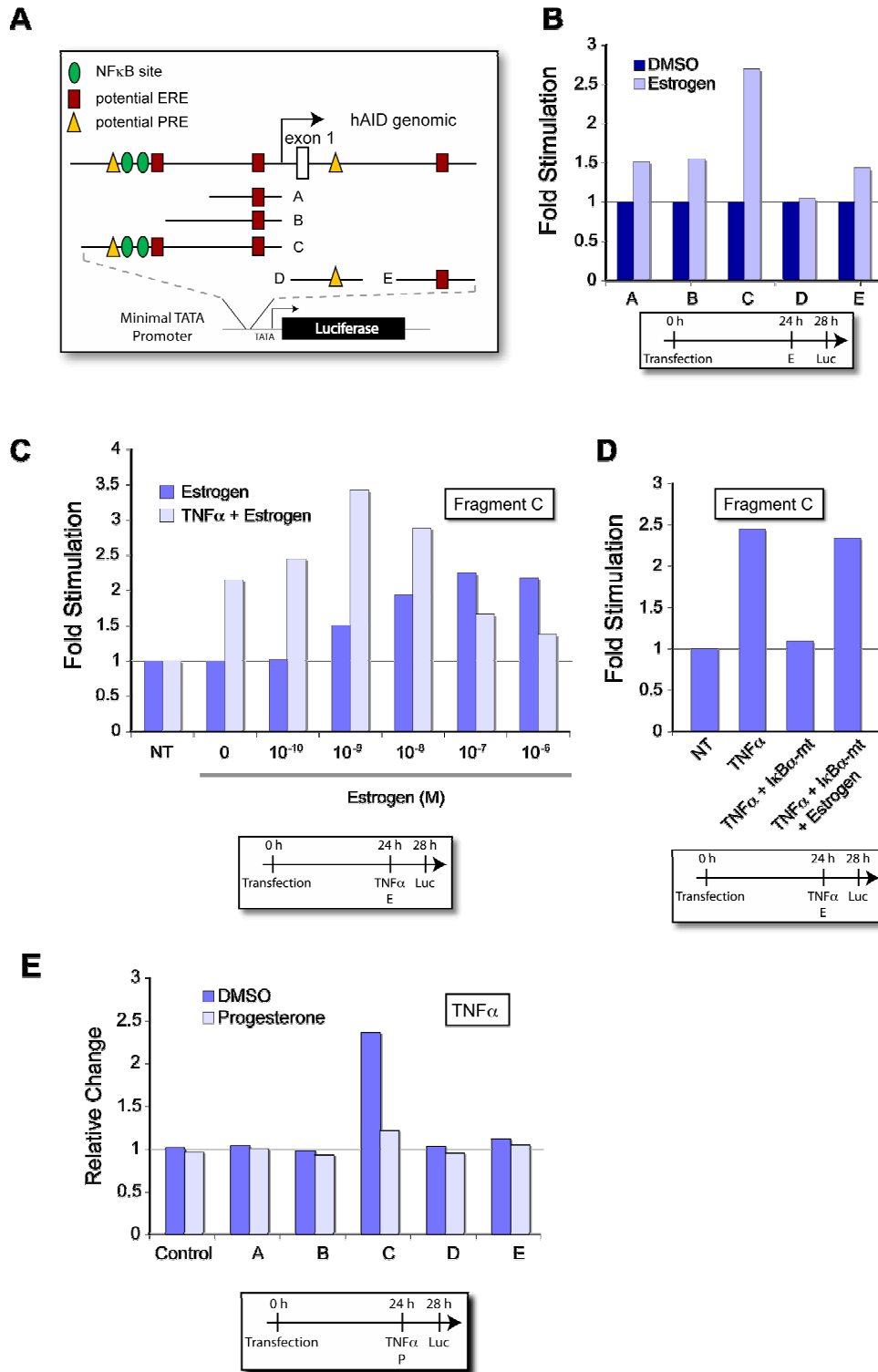
Figure 14. Human AID promoter analysis for hormone response elements. (A) Schematic representation of potential ERE (square), potential PRE (triangle) and NFκB sites (circle), and their locations in the human promoter. The indicated promoter regions (marked A-E) were inserted into a luciferase reporter construct with a minimal promoter. The vectors were transfected into SiHa cells, incubated for 24 h, treated for 4 h with hormones or TNFα and

analyzed for luciferase activity. (B) Relative luciferase activity after estrogen treatment. Cells were transfected with constructs containing AID promoter fragments and treated with estrogen for 4 h. (C) Effect of TNFα and estrogen on the human AID promoter. Expression construct with Fragment C (region containing NFκB sites and putative ERE and PRE) was transfected into cells followed by TNFα and/or estrogen treatment for 4 h. (D) Estrogen can act independently from NFκB. Cells were co-transfected with Fragment C and an IκBα dominant mutant expression vector. After 24 h, cells were treated with TNFα and/or 100 nM estrogen for 4 h. (E) Progesterone represses TNFα mediated induction of AID. Constructs were transfected into cells followed by TNFα treatment or TNFα and 200 nM progesterone co-treatment or for 4 h. Time-lines of cell treatments are indicated below the graphs. NT - not treated. Published in (Pauklin et al., 2009b; Pauklin and Petersen-Mahrt, 2009).

4.1.2.4 Estrogen response element binding in B-cell extracts.

Since we observed an increase in the luciferase signal in one of the AID promoter fragments containing a putative ERE, we determined if this element was bound by the estrogen receptor *in vitro*. We prepared nuclear extracts from estrogen treated and untreated Ramos cells (estrogen treatment on AID mRNA in Ramos cells is shown in Appendix Figure 2A) and analysed the binding of ER (focusing on the more widely expressed subtype ERα) to AID promoter segments by electromobility shift assay (EMSA) (Pauklin et al., 2009b). For this, a 34 bp fragment containing the 5' most putative ER binding site (Fragment C) was incubated with untreated and treated extract (Figure 17B lanes 2 and 3). Estrogen treatment induced a shift (arrow), which was not present when cells were treated with TNFα (lane 4). Treatment with estrogen and TNFα had the same effect as estrogen alone (lane 5), indicating that the two pathways act through different nuclear proteins. Competition experiments with an ER binding site (lanes 6 - 8) or a mutation of the proposed ER site (lanes 9 - 11), showed ER like binding kinetics. Using antibodies to the DNA binding domain of ERα strongly inhibited the formation of the estrogen induced band (Figure 17C lanes 2 vs. 3 - 5), but the shift was unaffected by a control antibody (lanes 2 vs. 6 - 8 arrow). The anti-ERα antibody induced the appearance of a high-molecular weight complex (triangle in lane 5), which could be due either to the supershift of a dimerized ERα or heterodimer ERα/ERβ, but this was not analysed further.
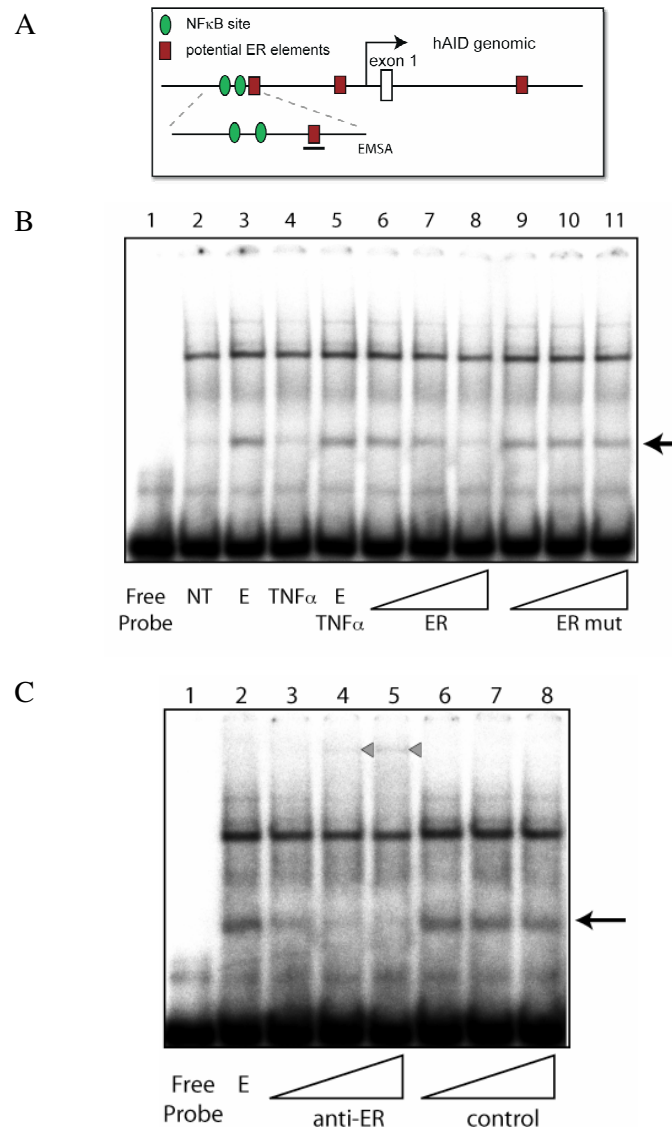
Figure 15. Identification of ER binding to human AID promoter by EMSA. (A) Schematic representation of human AID promoter region (as in Figure 16A). The position of the oligonucleotide used for EMSA is marked as a black line. (B) Estrogen (E) induced oligonucleotide shift (marked with an arrow) in Ramos nuclear extracts. Cells were treated for 72 h in hormone-depleted serum, followed by 4 h treatment with 10 nM estrogen (lanes 3, 5 - 11) and/or TNFα (lanes 4 and 5), and nuclear extract preparation. Different concentrations (1-, 3- and 10-fold excess) of unlabeled competitors ER (lanes 6 – 8) and mutated ER mut (lanes 9 - 11) were added to the binding reaction. (C) Supershift assays with anti-ERα antibodies. Increasing concentrations of anti-ERα antibody and a non-specific antibody were added to the binding reaction (see 3.2.2.22). The estrogen-induced band is marked with an arrow, and a super-shifted band appearing upon anti-ERα antibody addition is marked with triangles. Published in (Pauklin et al., 2009b).

In order to test, whether NFκB binding to the NFκB site was affected by estrogen after the combined treatment, we probed the published distal NFκB site

(Dedeoglu et al., 2004) in our extracts. Using cold competitors (Figure 18B lanes 4 vs. 6-11) as well as anti-NFκB antibodies (Figure 18C), the specificity of NFκB binding sites for binding of NFκB was shown. The specific antibody caused a supershift and the cold competitors induced a decrease in band intensity. As with the ER binding site, NFκB binding was not altered by co-treatment with estrogen and TNFα (Figure 18B, lanes 4 vs. 5). Thus, the respective treatments did not alter the general binding properties of ER or NFκB to the DNA. Since the distance on the AID promoter between the ER and NFκB sites was larger than the EMSA probes (i.e. neither probe contained both binding sites), cooperative binding of these transcription factors was studied by ChIP.
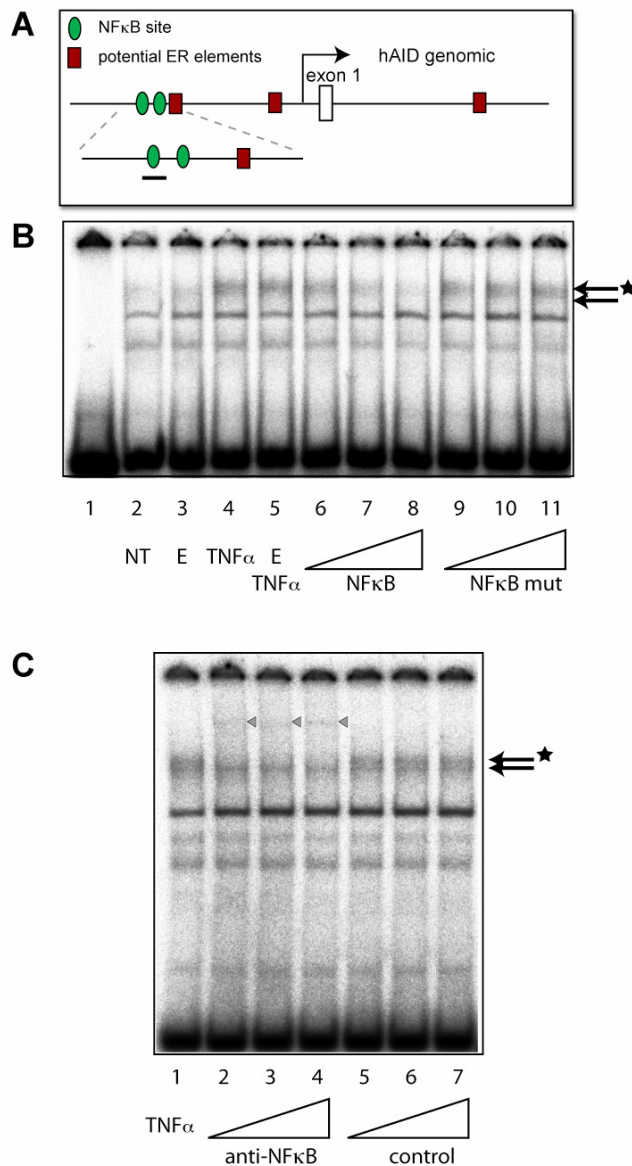
Figure 16. Estrogen treatment does not affect NFκB binding to its response element in human AID promoter. (A) Schematic representation of human AID promoter region as in Figure 16A. The position of the oligonucleotide used for EMSA containing the distal NFκB response element is depicted as a black line. (B) The appearance of two oligonucleotide shifts (marked with arrows) upon cell treatment with TNFα. Ramos cells were treated as in Figure 17. Different concentrations of cold competitors (NFκB - lanes 6-8) and mutated (NFκBmut - lanes 9-11) competitors were used to assess the specificity of these bands for NFκB binding. (C) EMSA with anti-NFκB antibodies. Cells were treated as in Figure 17 using anti-NFκB antibodies. The disappearing band and a super-shifted band, which appears upon anti-NFκB p65 antibody addition, are marked with a black arrow/star and triangles, respectively, indicating that one of the shifted bands contained NFκB p65. NT - not treated. Published in (Pauklin et al., 2009b).

4.1.2.5 Estrogen receptor and progesterone receptor binding to AID promoter *in vivo.*

In order to determine whether ER could bind to AID promoter *in vivo*, chromatin immunoprecipitation (ChIP) assays were performed on Ramos cells, followed by qRT-PCR. Treated cells were fixed, lysed, sonicated, ERα or NFκB immunoprecipitated, and the DNA released for PCR. The results indicated that the anti-ER antibody specifically immunoprecipitated AID promoter upon estrogen stimulation in a dose dependent manner, whereas a control antibody was unable to precipitate this region (Figure 19B). As in EMSA assay, an increase in ER binding to AID promoter was observed when cells were treated with estrogen alone or co-stimulated with TNFα and estrogen. The ERα binding to the AID promoter was partially abolished by co-treatment with tamoxifen (Figure 19B). Treatment of the cells with TNFα increased the binding of NFκB to the AID promoter (Figure 19C). Interestingly, co-stimulation of the cells with TNFα and estrogen had a synergistic effect on the NFκB binding - the binding increased 4.5 fold above TNFα alone. This suggests that NFκB and estrogen pathways can co-operate in the regulation of AID expression.

A

B

C
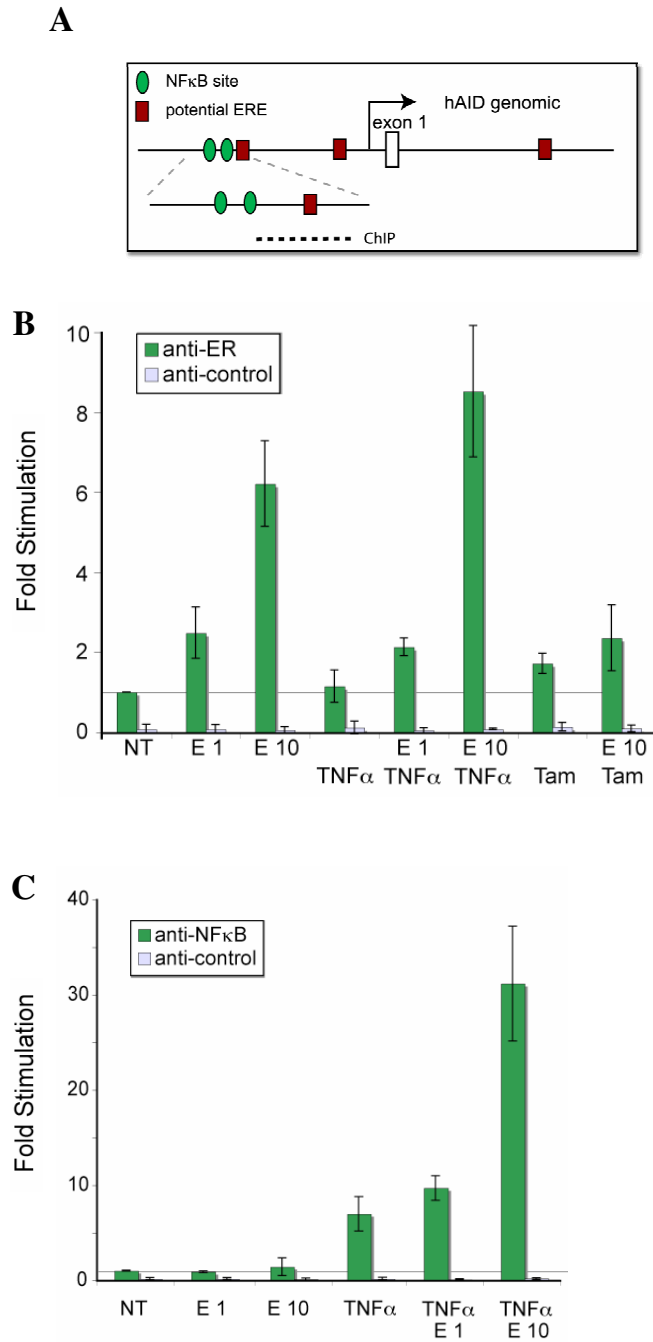
Figure 17. Identification of ER binding to human AID promoter by ChIP. (A) Schematic representation of human AID promoter region (as in Figure 16A). The region amplified by qRT-PCR for ChIP is marked as a dashed line. (B) ERα binds to upstream region of human AID promoter. Cells were treated as in Figure 17B and data is represented as a relative change compared to control (NT). ChIP was performed using anti-ERα or control antibodies and the bound DNA was subjected to qRT-PCR. Estrogen treatment is marked with E1 (E 1 nM) and E10 (E 10 nM). (C) Estrogen can cooperate with TNFα in recruiting NFκB to AID promoter. ChIP as in (B) using anti-NFκB or control antibodies. NT - not treated. Published in (Pauklin et al., 2009b).

To investigate whether progesterone exerts its effects via PR binding to AID promoter, Ramos cells were utilized for ChIP. Cells were treated with progesterone (the effects of progesterone on AID mRNA in Ramos are shown in Appendix Figure 2C), TNFα or with progesterone and TNFα simultaneously. Treatment of the cells with TNFα increased the binding of NFκB to the AID promoter (Figure 20E). The anti-PR antibody specifically immunoprecipitated the AID promoter upon progesterone treatment, but a control antibody failed to precipitate this region (Figure 20D). Interestingly, co-treatment of the cells with TNFα and progesterone revealed that both NFκB and progesterone receptor bound to AID promoter at the same time (Figure 20D and E). This suggests a mechanism, where progesterone receptor binding to the AID promoter, which leads to a decrease in AID transcription, does not abolish NFκB binding to the promoter. Since gene repression can be mediated by physical alteration of the interactions between transcription factors and the basal transcription machinery, one can not rule out the possibility that progesterone receptor mediated repression of AID promoter utilizes this mechanism. Primers near the NFκB sites in human AID promoter were used for qRT-PCR on ChIP samples. By comparing the relative signal increase that was amplified by these primers, the major binding region of PR was located to the proximity of the NFκB binding region (Figure 20B). It remains to be determined, whether progesterone receptors might also engage in the regulation of AID expression upon the activation of different signaling pathways, which rely on transcription activators other than NFκB.
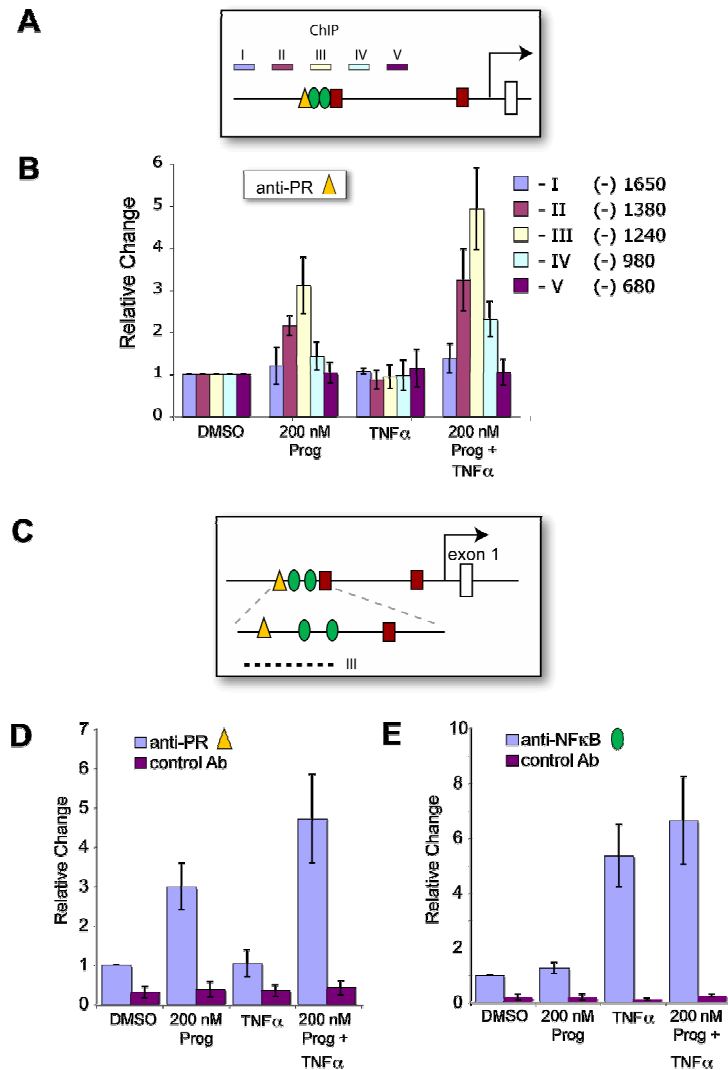
Figure 18. Progesterone receptor is recruited to the AID promoter. (A) Schematic drawing of AID promoter region. The five qRT-PCR products tested in the ChIP are indicated as (I-V). Triangle - putative PRE, rectangle - putative ERE and circles - NFκB sites (B) ChIP of 4 h un-, TNFα-, progesterone-, and TNFα/progesterone-treated Ramos cells. ChIP was carried out with anti-PR antibody, and the de-crosslinked DNA was analysed by qRT-PCR. Each PCR fragment was 150 bp in length and the center position (in bp upstream of the start site) of each of the fragments is indicated next to the color legend. Results are normalized to input DNA. Signal strength is presented as in Figure 19. Standard deviations are derived from three independent experiments. The time line of treatments is shown at the bottom of the figure. (C) Schematic drawing of AID locus with emphasis on ChIP region III (from A). (D) ChIP with anti-PR and control antibody followed by qRT-PCR of region III. Cells were treated or for 4 h; data analyzed as in B. (E) As D but using an anti-NFκB antibody for ChIP. Published in (Pauklin and Petersen-Mahrt, 2009).

4.1.2.6 Estrogen up-regulates whereas progesterone down-regulates AID protein production.

In order to determine if the effect of estrogen and progesterone on AID expression would also extend to the protein level, a quantitative approach for measuring AID protein was developed. For this, a DT40 cell line (a chicken B-cell lymphoma that constitutively expresses AID and undergoes immunoglobulin diversification; described in 4.2.1) was generated to express a 3FLAG-2TEV-3Myc fused to the C terminal exon of endogenous AID. The modified DT40 expresses a wt AID protein as well as an AID-FLAG-TEV-Myc fusion protein, both transcribed from the endogenous AID locus (the generation of the cell line is described in more detail in 3.2.3 and in 4.2.2.1).

By comparing the qRT-PCR induction kinetics with that of the AID-FLAG-TEV-Myc protein signal on a quantitative western blot, AID protein levels were found to correspond to mRNA levels when cells were treated with estrogen (Figure 21A and B). By using the translation inhibitor cycloheximide (Figure 21C) we were able to assess AID protein degradation as published previously (Aoufouchi et al., 2008). Estrogen treatment did not increase tagged AID protein sfter cycloheximide treatment, indicating that this hormone did not affect AID protein stability but functioned at the gene transcription stage. On the other hand, proteasome inhibitor MG-132 treatment increased the fusion protein levels compared to MG-132 alone (Figure 21C), consistent with the observations by (Aoufouchi et al., 2008). In addition, estrogen treatment resulted in a further increase in AID fusion protein at a comparable change as estrogen treatment alone, suggesting that estrogen was not affecting the proteasome to increase AID-FLAG-TEV-Myc protein.
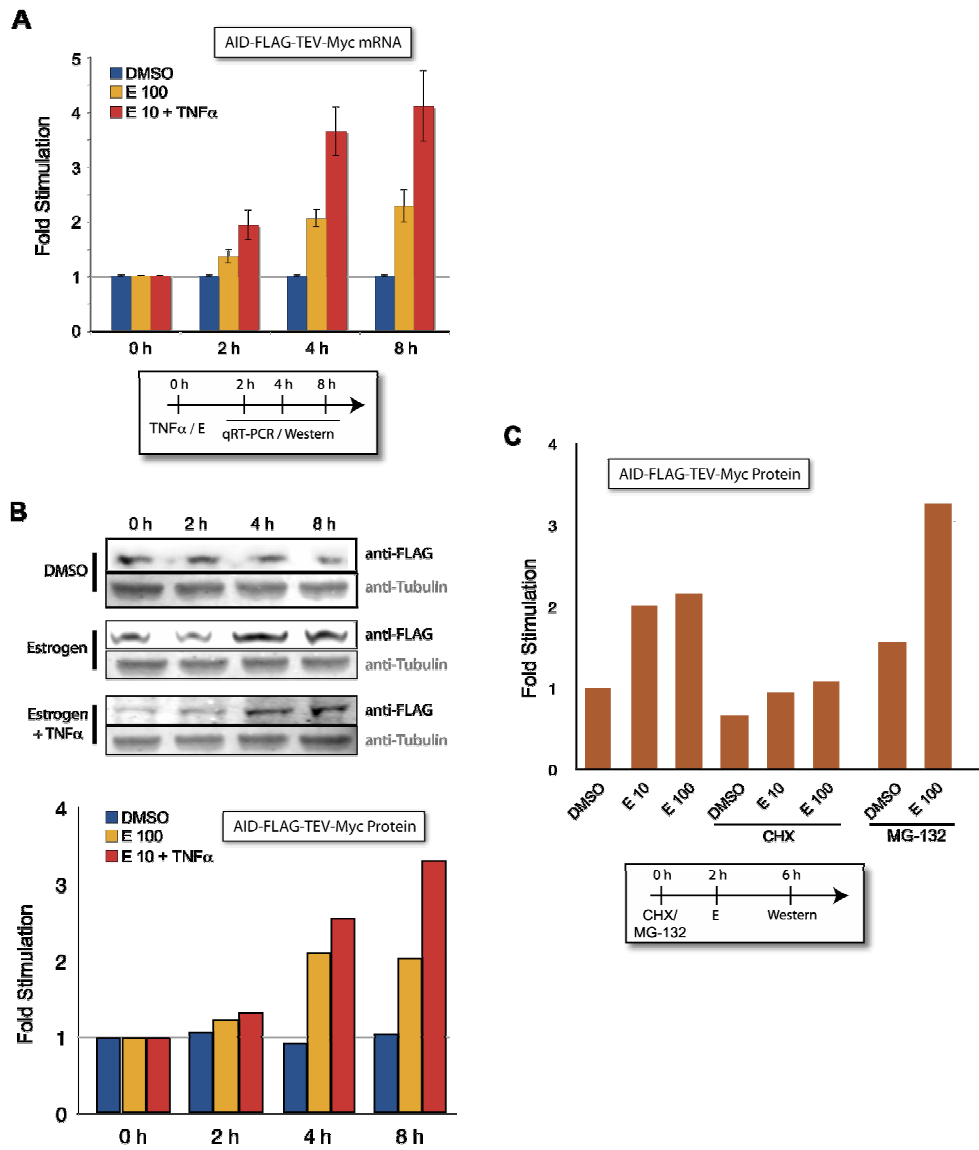
Figure 19. The effects of estrogen on AID protein in DT40. (A) Estrogen induces AID mRNA expression. DT40 cells tagged with AID-FLAG-TEV-Myc were treated with DMSO, 100 nM estrogen, and 10 nM estrogen with TNFα, and analyzed for AID mRNA expression with qRT-PCR at various time-points. (B) Estrogen induces AID-FLAG-TEV-Myc fusion protein expression. Treatment as in (A) but lysates were analyzed by quantitative Western blot (see 3.2.2.18). For each sample, FLAG and Tubulin expression was quantitated. The bar graph is derived from correlating the FLAG expression to Tubulin expression, and then determining the ratio of estrogen induced FLAG expression to untreated DMSO samples. (C) Estrogen does not affect AID-FLAG-TEV-Myc fusion protein stability. Cells were incubated with CHX or MG-132 for 2 h, prior to estrogen treatment for 4 h. Protein levels were determined by quantitative Western blot. For all experiments, cells were grown in hormone depleted media for 48 h. Results are normalized to control treatments as indicated on each graph. Time-lines of cell treatments are indicated below the graphs. Published in (Pauklin et al., 2009).

Cells were also treated with TNFα or TNFα together with progesterone and analysed for mRNA and protein levels of AID in response to these treatments. As

with the effect on mRNA (Figure 22A), progesterone treatment of the modified DT40 cells showed a decrease in AID-FLAG-TEV-Myc production over time (Figure 22B). This indicated that progesterone exerted its effects on AID mRNA and thereby on AID protein.
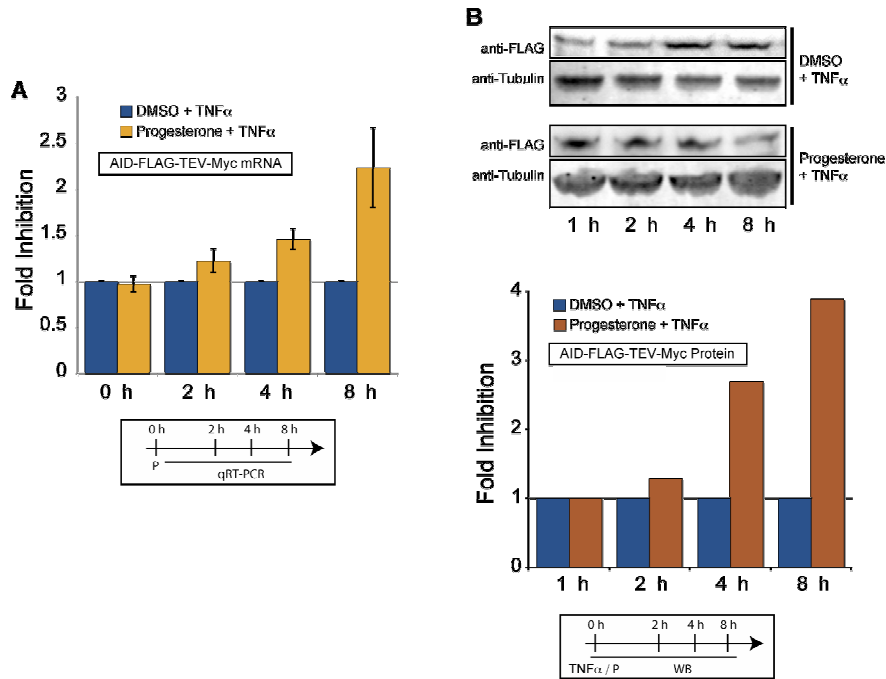
Figure 20. The effects of progesterone on AID protein in DT40 cells. (A) Progesterone inhibits AID mRNA expression in DT40. AID-FLAG-TEV-Myc tagged DT40 cells were treated with 200 nM progesterone and/or TNFα, and then analyzed for AID mRNA expression by qRT-PCR at various time-points (3 samples for each treatment). (B) Progesterone inhibits tagged AID protein expression. Treatment as in (B) but lysates were analyzed for FLAG epitope expression using quantitative Western (pooled analysis from 3 samples per treatment). Images of Western are shown above, and quantitations of bands below. Time-line of treatments is indicated below the figures. Published in (Pauklin and Petersen-Mahrt, 2009).

4.1.2.7 Hormonal regulation of CSR, SHM, and translocations.

AID is essential for triggering class switch recombination, therefore the potential effects of estrogen and progesterone mediated regulation of AID were analysed on the level of class switch recombination. B-cell stimulation with LPS, LPS and IL4, LPS and IFN-γ, or LPS and TGF-β causes isotype switching to IgG3, IgG1 and IgE, IgG2a or IgG2b/ IgA, respectively. During this process, a circular DNA called a switch circle is formed, that contains a recombined transcription unit, producing a switch circle transcript (Appendix Figure 5). Using circle transcript analysis combined with qRT-PCR (Kinoshita et al., 2001), alterations were

observed in the switching to IgG1, IgG3, IgA or IgE after hormone treatment of splenic B-cells (Figure 23A). Similarly to AID mRNA induction, estrogen enhanced switch circle formation and this increase was attenuated by tamoxifen, although additional effects of estrogen on circle transcript formation can not be ruled out. In addition to switch circle analysis, we sequenced the $\gamma3$ switch region in AID heterozygous mice stimulated with LPS for 6 days and found an increase in mutations in this locus in response to estrogen treatment compared to the control sample (Figure 23B; Appendix Figure 3C).

To further analyse SHM, a derivative (Ramos HS13) of the constitutive AID expressing and mutating Burkitt lymphoma Ramos was utilized (Sale and Neuberger, 1998). This cell line possesses a premature stop codon in the Ig λ locus within an AID target motif WRC. When AID targets this codon, it can be converted to a sense codon, allowing IgM surface expression, thereby giving a possibility to analyse SHM. In Ramos HS13 cells, AID mRNA was modestly up-regulated upon estrogen treatment (Appendix Figure 2A). Surface IgM negative cells were single cell sorted and grown in the presence of hormones for 20 cell generations (3 weeks), followed by surface IgM expression analysis (Appendix Figure 2B). In a similar trend to AID mRNA, surface IgM expression was increased by estrogen treatment. To monitor a change in mutation frequency, the VH region of estrogen treated cells was also sequenced, which showed an increase in mutation frequencies (Figure 23B - 'Ramos VH').

AID can be mis-targeted to non-immunoglobulin genes, causing mutations and subsequently oncogenesis. To determine, if estrogen induced AID can also lead to an alteration in non-Ig loci, the sequence of the pro-apoptotic tumor-suppressor CD95/Fas was examined in Ramos HS13. This gene has been shown to be somatically hypermutated in human B-cells, albeit 100 to 1000 fold less frequently than the Ig genes (Muschen et al., 2000). The genomic DNA of the above treated cells were PCR amplified and sequenced in the 5' region of the Fas locus. Similarly to the B-cell maturation events of SHM and CSR, estrogen was able to increase the mutation frequency in Fas (Figure 23B). Because of the direct effect of AID function on mutations, the upregulation of AID expression by estrogen indicates a novel way in which estrogen can exert a direct mutagenic effect on oncogenes or tumor suppressors.
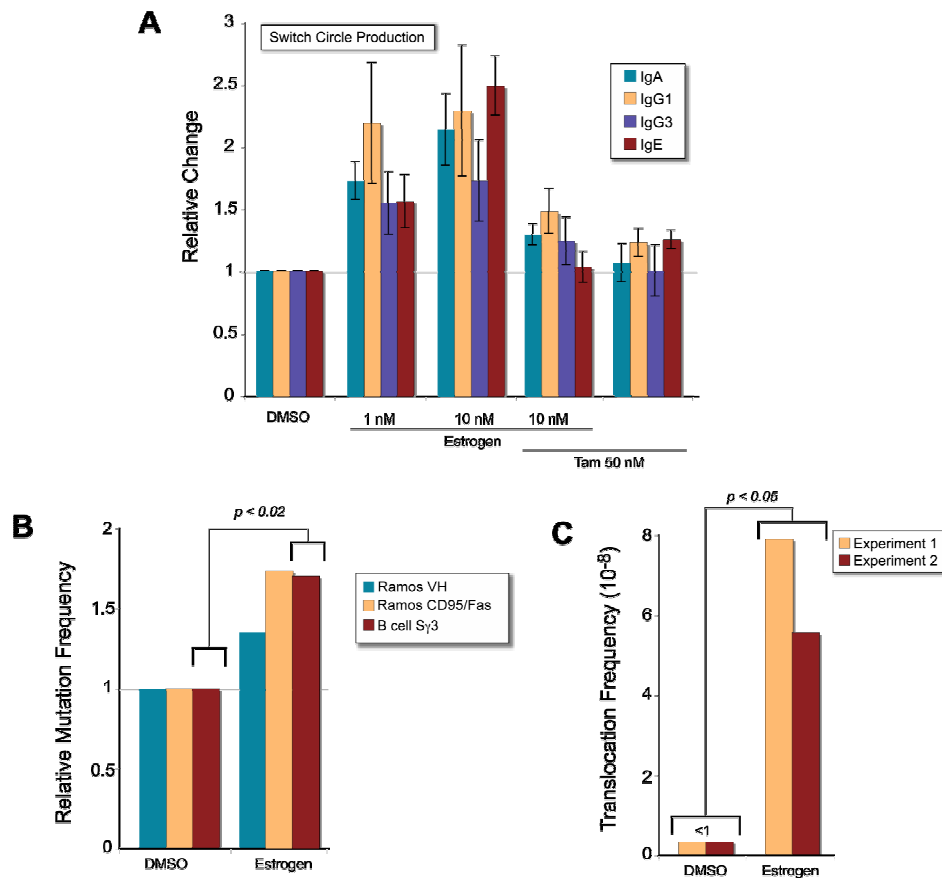
Figure 21. Hormonal effects on Ig class switching, hypermutation, and translocation. (A) Estrogen induces Ig class switching. Isolated mouse splenic B-cells were stimulated for 48 h with LPS + IL4 for switching to IgG1 and IgE, LPS + TGF-β for switching to IgA, and LPS for switching to IgG3. Indicated amounts of estrogen and/or tamoxifen were added to the cells together with cytokines. Relative efficiency of class switching was determined by detecting circle transcripts with qRT-PCR and data is normalized to the control treatment with DMSO from three independent experiments (error bars indicate standard deviations). (B) Estrogen increases the mutation frequency in VH and CD95/Fas loci of Ramos, and Sγ3 of splenic mouse cells. Ramos cells were grown in the presence of 100 nM estrogen for approximately 20 doublings, followed by sequencing of 341 bp from human VH, or 750 bp from human CD95/Fas locus. Splenic mouse cells were treated for 6 days with LPS and 10 nM estrogen (details in Appendix Figure 3), and switch Sγ3 loci amplified and sequenced. Mutation frequencies are normalized to the control treatments with DMSO. A standard unpaired two-tailed T-test showed a significant difference in mutation frequency in the Sγ3 loci of DMSO and estrogen treated spleen cells. (C) Estrogen enhances the c-myc/IgH translocations in splenic B-cells from p53 +/- mice. In each experiment 2 spleens per sample were treated with or without 50 nM estrogen in the presence of LPS for 72 h. More than 7×10^7 cells were analysed by long-range PCR (5×10^4 cells/PCR; see 3.2.2.16). Frequency was determined as c-myc/IgH translocation events per cell number analysed. Statistics was performed on the results of the pooled experiments (two tailed, unpaired T-test: $p = 0.026$). Published in (Pauklin et al., 2009b).

The off-target effect of AID can also be observed in translocation events between the IgH locus and c-myc oncogene (Ramiro et al., 2004). Using a previously described assay, we found that estrogen treatment of LPS stimulated splenic B-cells increased c-myc/IgH translocations in p53 +/- animals (Figure 23C). Despite a low frequency of translocation of events, we found five (2 in one and 3 in another experiment) translocations in the LPS/estrogen treatment, whereas no translocations were observed in the control LPS/DMSO treated sample. This suggests, that mis-targeting of estrogen-induced AID can cause translocations in thus have pathogenic consequences to the cell.

As for estrogen, we also analysed the effects of progesterone on class switching and somatic hypermutation. Progesterone treatment of stimulated B-cells resulted in a decrease in class switching (Figure 24A), mimicking the attenuation of AID mRNA induction upon B-cell stimulation (Figure 11A). Surface IgM analysis in Ramos HS13 cells indicated a decrease in SHM
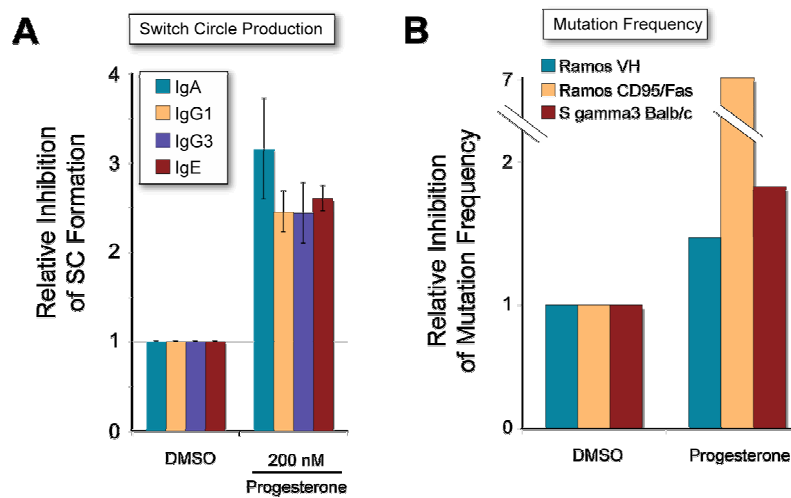
Figure 24. Progesterone inhibits CSR and SHM. (A) Progesterone inhibits isotype switching. Isolated mouse splenic B-cells were stimulated for 48 h with LPS + IL-4 for switching to IgG1 and IgE, LPS + TGF-β for switching to IgA, and LPS for switching to IgG3. Indicated amounts of progesterone were added to the cells together with cytokines. Relative efficiency of class switching was determined by detecting circle transcripts with qRT-PCR and data is normalized to the control treatment with DMSO from three independent experiments (error bars indicate standard deviations). (B) Progesterone decreases the mutation frequency. Using a cell culture surface IgM reversion assay, human Ramos cells were grown as indicated in Appendix Figure 4A and B, followed by sequencing of 341 bp from the rearranged VH (VH186) or 750 bp from human CD95/Fas locus. Mouse splenic B-cells were treated for 6 days with LPS and 200 nM progesterone. Genomic DNA was amplified and 750 bp of the Sγ3 region sequenced (details provided in Appendix Figure 4C). Mutation frequencies are normalized to the control treatments with DMSO. Published in (Pauklin and Petersen-Mahrt, 2009).

(Appendix Figure 2D and E). In addition, there was a decrease in mutation frequencies in the VH region and CD95/Fas locus in Ramos HS13, and a decrease in mutations in γ3 switch region in mouse splenic B-cells (Figure 24B and Appendix Figure 4).

4.1.2.8 AID induction by estrogen is not limited to B-cells.

AID expression has been detected in tissues outside B-cells such as oocytes (Morgan et al., 2004). We therefore analysed whether the increase in AID mRNA production by estrogen would also be detectable in dissected tissues. The analysis showed, significant AID responsiveness to hormones in ovarian tissue and to a lesser extent in breast tissue, even though the signal was much lower compared to B-cells (Figure 25 and 26). Since the relative increase in AID mRNA was higher
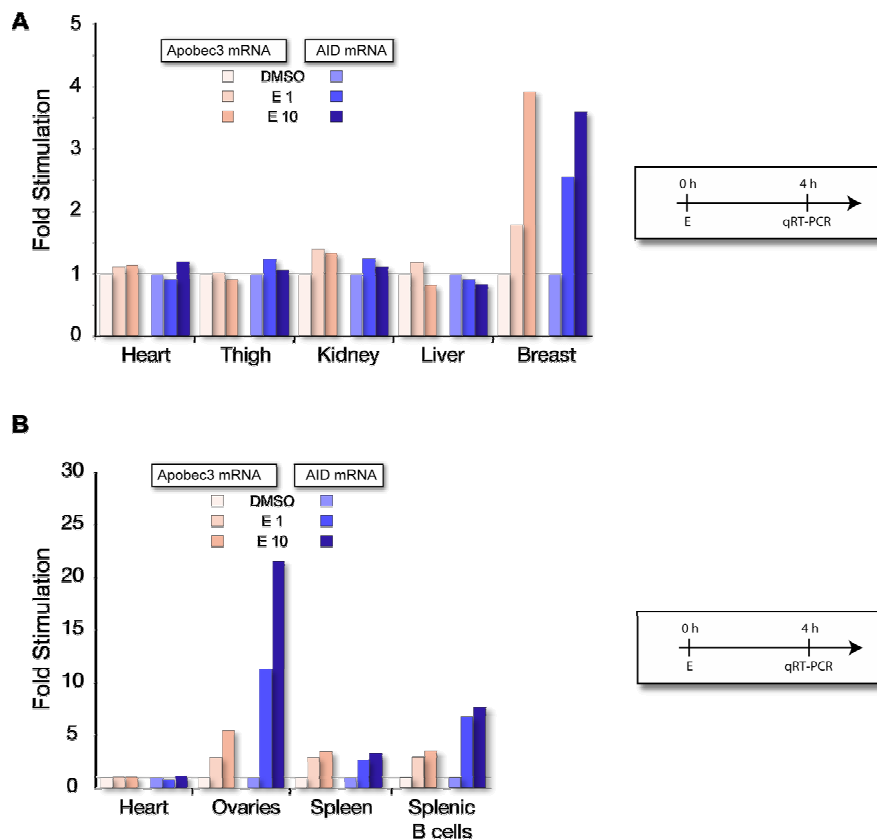
Figure 25. Estrogen induces AID and Apobec3 transcription in mouse tissue. The red and blue colors indicate the results for mApobec3 and mAID, respectively. Tissues were treated with DMSO, 1 nM estrogen (E1), or 10 nM estrogen (E10). Gene expression is normalized to the control treatments with DMSO. The tissue expression profiles represent pooled data for the respective tissues from two experiments. Time-lines of cell treatments are indicated below the graphs. Published in (Pauklin et al., 2009b).

in oocytes (reaching approximately 25 fold) than in other tissues, it is likely that this increase was predominantly due to ovarian derived tissue and not contaminating B-cells.
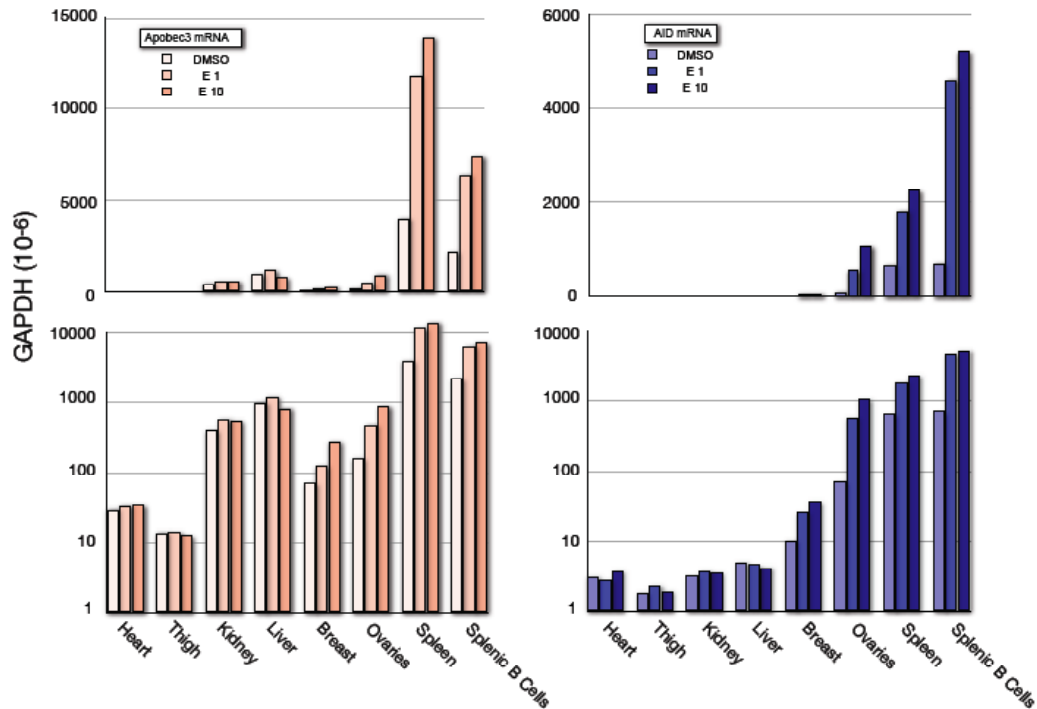
Figure 22. Absolute values of AID and Apobec3 mRNA in mouse tissues. Absolute qRT-PCR values as compared to GAPDH expression from Apobec3 data (light red) and AID data (blue) of Figure 25. In the upper panels, values are plotted on a linear scale, and in the lower panels, values are plotted on a logarithmic scale. Published in (Pauklin et al., 2009b).

4.1.2.9 Estrogen activates APOBEC3B, 3F and 3G mRNA transcription.

The regulation of AID by estrogen raised the question whether a similar regulation is conserved among other AID/APOBEC family members. qRT-PCR primers to detect APOBEC family member mRNAs in mouse and human cells were designed and the expression of APOBEC genes examined upon estrogen treatment (Figure 25 and 26). Estrogen treatment led to an increase in transcription of the mouse Apobec3 in ovaries, spleen, and splenic B-cells. Expression of human APOBEC3B, 3F and 3G family members were also enhanced upon estrogen treatment in several different cell lines (including: T-cell, ovarian, placental, and cervical origin) (Table 10). APOBEC2, a related

member of the DNA deaminases without any apparent catalytic activity (Conticello et al., 2005), was not affected by estrogen treatment. Since AID is considered as the ancestor to APOBEC3 genes, the regulation of estrogen of AID and other family members suggests that this might be conserved mechanism among these genes.

Table 10. Estrogen induces the transcription of AID and some APOBEC3 family members in various hormone-responsive human cell lines. The indicated human cell lines were treated with 1 nM or 10 nM estrogen (E1, E10) or 50 nM tamoxifen (T50) for 4 h, followed by gene expression analysis by qRT-PCR. Results are normalized to control treated cells. Small numbers below the fold induction indicate standard deviations for three independent experiments. Yellow, orange and red rectangles indicate the extent of fold induction in response to these treatments. Time-line of cell treatments are indicated below the table. Published in (Pauklin et al., 2009b).

	Jurkat (T cell)			JAR (Placenta)			JAMA2 (Ovarian)			MCF7 (Breast)			HeLa (Cervix)			PC3 (Prostate)			HepG2 (Liver)			T47D (Breast)		
	E 1	E 10	T 50	E 1	E 10	T 50	E 1	E 10	T 50	E 1	E 10	T 50	E 1	E 10	T 50	E 1	E 10	T 50	E 1	E 10	T 50	E 1	E 10	T 50
AID	10.1	3.19	2.19	5.81	1.54	2.58	1.36	4.78	2.92	7.64	2.80	1.69	0.97	1.35	1.24	6.92	21.8	4.93	1.66	0.76	1.62	4.42	2.10	1.00
	1.93	0.71	0.54	0.94	0.15	0.96	0.17	0.81	0.36	0.63	1.8	0.13	0.34	0.48	0.42	1.43	5.35	0.96	0.4	0.9	0.32	0.7	0.63	0.3
APOBEC2	0.96	1.26	0.99	1.30	1.13	1.00	1.30	1.15	0.77	1.39	1.60	0.88	1.19	1.50	0.94	1.21	1.24	1.40	1.16	1.26	0.90	1.15	1.28	1.00
	0.15	0.17	0.14	0.11	0.09	0.09	0.49	0.38	0.16	0.25	0.22	0.14	0.33	0.33	0.25	0.4	0.28	0.19	0.32	0.17	0.14	0.38	0.2	0.2
APOBEC3A	2.30	2.36	1.38	1.79	2.05	1.46	1.05	0.87	1.14	1.52	1.85	0.08	1.3	1.99	1.25	1.53	1.95	1.18	0.78	0.79	0.99	1.25	1.28	1.09
	0.88	0.61	0.37	0.11	0.11	0.38	0.08	0.10	0.32	1.18	0.63	0.35	0.36	0.41	0.34	0.23	0.20	0.33	0.03	0.06	0.23	0.27	0.06	0.24
APOBEC3B	2.46	3.90	1.69	1.92	2.63	1.31	1.75	2.42	1.47	1.26	1.83	1.50	2.31	4.89	1.53	1.39	1.31	0.87	1.26	1.45	1.24	1.18	1.41	1.35
	0.35	0.76	0.39	0.26	0.38	0.16	0.22	0.61	0.29	0.13	0.44	0.15	0.29	0.94	0.14	0.46	0.27	0.19	0.3	0.33	0.12	0.27	0.35	0.7
APOBEC3C	1.75	1.89	1.36	1.91	3.28	1.48	1.22	1.07	0.94	1.33	1.32	1.01	0.88	1.29	1.1	1.54	1.5	1.34	0.94	0.85	0.85	1.23	1.71	1.28
	0.33	0.39	0.28	0.17	0.36	0.28	0.20	0.21	0.34	0.17	0.11	0.37	0.06	0.19	0.27	0.04	0.30	0.15	0.26	0.19	0.07	0.09	0.44	0.32
APOBEC3D	1.10	1.31	0.97	1.24	1.56	0.95	1.1	1.41	1.48	1.55	1.54	0.87	1.21	1.44	1.08	1.95	3.61	2.15	1.39	0.98	0.81	1.31	3.51	1.48
	0.16	0.22	0.23	0.35	0.28	0.02	0.47	0.22	0.18	0.18	0.20	0.11	0.29	0.25	0.02	0.32	0.35	0.35	0.18	0.16	0.12	0.30	0.08	0.37
APOBEC3F	1.94	4.35	0.97	2.20	4.46	1.96	1.62	4.20	1.98	1.35	1.89	0.96	1.77	6.56	2.35	1.46	0.90	1.00	1.19	1.38	1.10	0.99	1.97	1.80
	0.13	0.7	0.2	0.51	0.8	0.57	0.63	1.55	0.81	0.15	0.65	0.19	0.45	0.87	0.33	0.3	0.13	0.19	0.5	0.45	0.16	0.7	0.62	0.23
APOBEC3G	2.57	3.74	1.46	2.20	3.14	1.45	1.58	2.22	1.68	1.55	1.60	0.98	2.80	5.19	2.30	1.58	1.47	1.20	1.60	1.35	1.23	1.27	1.97	1.10
	0.23	1.5	0.21	0.31	0.31	0.09	0.39	0.78	0.22	0.57	0.05	0.06	0.38	1.27	0.32	0.29	0.44	0.17	0.29	0.43	0.06	0.05	0.62	0.22
APOBEC3H	1.98	2.92	1.85	1.83	2.56	1.27	1.16	0.96	1.62	1.4	1.57	1.42	0.98	1.32	1.42	1.42	1.55	1.18	1.34	1.56	0.68	1.12	1.69	1.02
	0.24	0.33	0.22	0.11	0.30	0.18	0.11	0.08	0.35	0.20	0.11	0.27	0.11	0.08	0.14	0.10	0.22	0.09	0.27	0.20	0.10	0.06	0.28	0.13

fold increase in signal

2.5 to 4.5	4.5 to 6.5	> 6.5

0 h — 4 h, E — qRT-PCR
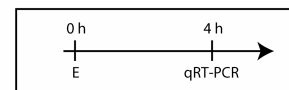
4.1.3 Discussion

The inactivation of AID abolishes SHM and CSR, as has been observed in AID knockout mice (Muramatsu et al., 2000) and human patients with the type 2 hyper IgM (HIGM) syndrome (Revy et al., 2000). On the other hand, if overexpressed or misregulated, AID can act as a mutator and lead to translocations and mutations in several oncogenes and tumorsuppressors. Therefore, AID expression needs to be regulated, to allow for efficient immune responses, and at the same time, to avoid the harmful genotoxic effects. In this study, I have identified steroid hormones estrogen and progesterone as regulators of AID. The regulation of AID by estrogen and progesterone seems to be direct, since estrogen receptor and progesterone receptor are able to bind AID promoter (Figure 19 and 20). Estrogen is able to increase the transcription of AID, thereby resulting in an upregulation of AID mRNA (Figure 11A) and protein (Figure 21B). Progesterone, on the other hand, decreases AID mRNA (Figure 11A) and protein production (Figure 22B). In addition, these ER and PR engage in crosstalk with factors that participate in AID regulation, including the NFκB pathway (Figure 16). The interaction between estrogen and NFκB is complex, since the co-activation of these factors can result in synergistic induction of AID expression at low estrogen concentrations, but looses this effect at highes estrogen concentrations (Figure 16C). On the other hand, progesterone represses NFκB mediated induction of AID expression (Figure 16E).

Although hormones are involved in many aspects of development, the extent of their functioning, as well as their exact molecular effector mechanisms, are not well understood. There is a correlation between the expression of hormones or their cognitive receptors and the development of some gender related cancers and autoimmune diseases. Estrogen and its biological and synthetic derivates are considered oncogenic for at least breast and ovarian tissue (Roy and Liehr, 1999; Yager and Davidson, 2006). Early menarche, late menopause and length of reproductive life, and factors consistent with prolonged exposure to endogenous estrogens, have consistently been associated with an increased risk of breast cancer in post-menopausal women. This is at least partially caused by the stimulation of cell proliferation, changes in differentiation and DNA repair and repression of apoptosis (Yager and Davidson, 2006).

However, my results suggest an additional possibility how estrogen might promote cancer – that is via upregulating AID expression. These results indicate that AID expression is increased by estrogen (Figure 11A), which casuses stimulation of its downstream processes, somatic hypermutation (Figure 23B) and class switch recombination (Figure 23A). Besides the Ig locus, there seems to be an increase in AID mediated mutations also in non-Ig loci, since the mutation frequency in tumorsuppressor Fas was increased upon estrogen treatment (Figure 23B). This imposes a threat to the cell that expresses AID, since mutations in tumor suppressors or oncogenes can result in the deregulation of apoptosis or constitutive cell growth, ultimately leading to malignancies. Indeed, previous work has shown that somatic hypermutation can result in mutations in various non-Ig loci such as *BCL-6, Fas, p53, or c-myc* (Kuppers and Dalla-Favera, 2001; Muschen et al., 2000; Pasqualucci et al., 2001). Furthermore, forced AID expression outside the immune system can cause tumor formation (Okazaki et al., 2003). This is relevant for ovarian tissue, since I observed approximately 25-fold induction of AID in mouse ovaries (Figure 25 and 26). The treatment of the cells with estrogen also led to an increase in c-myc translocations (Figure 23C). These translocations are frequently identified in lymphomas, but analogous translocations could potentially occur in some hormone-responsive tissues upon estrogen mediated induction of AID.

In addition to oncogenesis, defects in apoptosis are also attributed to some B-cell mediated autoimmune pathologies. Since estrogen is involved in the gender bias in some autoimmune diseases such as SLE, the regulation of AID by this hormone could mechanistically link gender related malignancies and autoimmune pathologies. Over-expression of AID in autoimmune prone mice induced a more severe SLE-like disease (Hsu et al., 2007), while breeding AID deficient mice with autoimmune prone MRL/lpr mice reduced the on-set and extent of disease (Jiang et al., 2007), suggesting that alterations in AID can change the severity of B-cell autoimmunity. Future studies could address the role of estrogen induced AID in mouse models of these diseases.

In contrast to the effects of estrogen on AID expression and downstream effects, progesterone represses AID function in somatic hypermutation (Figure 24B) (both in the Ig locus and CD95/Fas gene) as well as class switch recombination (Figure 24A). The attenuating effect of progesterone on AID-

induced processes could therefore decrease some immune responses, which might be relevant during pregnancy, helping to avoid the destruction of the fetus by mother's immune system (Druckmann and Druckmann, 2005). The decrease in mutation frequencies outside the Ig locus indicates, that this regulatory mechanism could in some instances result in fewer off-target pathological mutations due to progesterone.

Interestingly, the hormonal regulation of DNA deaminases was not limited to AID, but was also evident (to a varying extent) for murine Apobec3 and several human APOBEC3 family members, especially for APOBEC3B, -3G and -3F (Figure 25 and Table 10). These proteins are essential anti-viral factors in higher mammals, inhibiting the movement of retro-transposable elements and retroviral infections (Harris and Liddament, 2004). Therefore, it is possible that the upregulation of APOBEC3 genes by estrogen in some cell types, such as oocytes, might have a physiological role in inhibiting the movement of retro-transposable elements.

In conclusion, I have identified estrogen and progesterone as novel regulators for AID that influence AID expression and its down-stream processes that could contribute to the regulation of immune responses but also be involved in pathological conditions such as cancer and autoimmunity.

4.2 Identification of AID-interacting proteins.

4.2.1 Introduction.

When AID was discovered, it was proposed to work on RNA as an RNA-editing enzyme (Muramatsu et al., 2000; Muramatsu et al., 1999). This notion was indirectly supported when AID was identified to reside predominantly in the cytoplasm (Rada et al., 2002). However, subsequent studies have shown that AID can enter the nucleus and initiate Ig diversification by deaminating cytosines in ssDNA [Reviewed in (Di Noia and Neuberger, 2007; Petersen-Mahrt, 2005)]. Since accessibility of DNA to the enzymatic activity of AID can potentially result in harmful mutations outside the Ig locus, it is likely that sub-cellular localization is an important level for regulating AID protein *in vivo*. Even though the localization of AID has been addressed in some studies (Aoufouchi et al., 2008; Ito et al., 2004; McBride et al., 2004; Patenaude et al., 2009; Rada et al., 2002), it is essential to study the regulation of the sub-cellular distribution of AID protein at its endogenous levels in B-cells that accurately recapitulates *in vivo* conditions.

An important mechanism for regulating AID activity is via interacting with other proteins that influence AID's localization, post-translational modifications, and targeting to the V and switch regions. Even though the identification of AID-interacting proteins in an unbiased setting is largely missing, several AID-interacting proteins have already been found. AID can be phosphorylated by PKA and interact with RPA in a phosphorylation dependent manner (Basu et al., 2005; Chaudhuri et al., 2004). The nuclear import of AID seems to be mediated by its binding to importin α (Patenaude et al., 2009). Further potential AID-interacting proteins include, Mdm2 (MacDuff et al., 2006) and DNA-PKcs (Wu et al., 2005), both can bind to the C-terminus of AID protein. In addition, AID co-immunoprecipitates with RNA polymerase II (Nambu et al., 2003) and binds directly to CTNNBL1, a protein involved in RNA splicing (Conticello et al., 2008).

Once AID accesses the Ig locus, the distribution of AID-induced mutations in this gene is not uniform (Figure 27). The first 100-200 bp from the

transcription initiation site are spared from mutations, and most of the mutations occur after this initial region, followed by an exponential decrease during the next 1.5-2 kb (Both et al., 1990; Lebecque and Gearhart, 1990; Longerich et al., 2006; Rada and Milstein, 2001). The frequency of mutations is highest in the V(D)J region, but it is unknown what causes this pattern of mutations. Since AID can interact with RNA polymerase II (Nambu et al., 2003), it is possible that the transcription machinery may be involved in establishing this characteristic
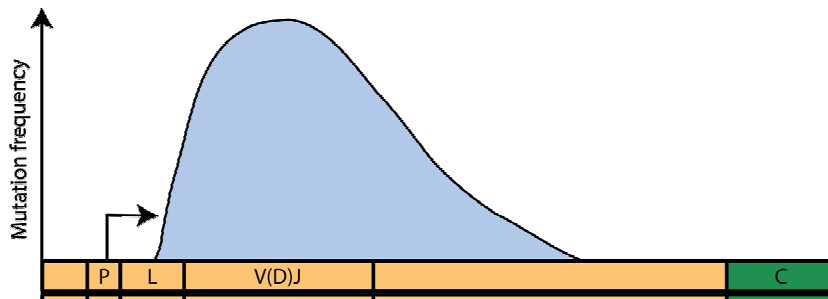
Figure 27. Schematic depiction of mutation distribution in the IgH locus. The variable region V(D)J is most heavily mutated whereas the constant region (C, green) is not mutated. SHM (blue) are delimited by the variable region promoter (P), appear 100-200 bp after transcription start site, peak in the variable region and exponentially decrease 1.5-2 kb 3' from the transcription start site. L marks leader sequence.

mutation distribution found in the Ig loci. The lack of mechanistic understanding of AID mediated DNA deamination comes from the absence of data investigating AID-interacting proteins that modulate DNA deamination *in vivo*. The identification of AID-interacting proteins is difficult due to the lack of a suitable system: 1) AID is expressed in activated B-cells, restricting the choice of cells, 2) AID protein is predominantly localized in the cytoplasm, making it difficult to identify proteins that form a complex with AID in the nucleoplasm and chromatin, and 3) there is a lack of antibodies that would immunoprecipitate AID protein and thus allow biochemical studies of AID in large enough quantities for complex analysis.

The avian DT40 cell line has several favourable characteristics (Buerstedde and Takeda, 2006), which make it a suitable system for studying the biochemistry of AID and for identifying its interaction partners: 1) it is a B-cell line that expresses endogenous AID; 2) it undergoes immunoglobulin diversification,

mainly via gene conversion, making it possible to analyse AID-mediated processes; 3) it has a high rate of homologous recombination, enabling the targeting of endogenous loci; and 4) it has a relatively short doubling time (10-12 hours), allowing for the growth of these cells in large quantities for protein interaction studies. These features of DT40 prompted us to generate DT40 cell lines with tagged AID, to study the localization of endogenous AID and identify AID-interacting proteins in an unbiased manner.

4.2.2 Results.

4.2.2.1 Generation of endogenously tagged DT40 cell lines.

4.2.2.1.1 Generating Tagged AID Constructs.

In order to generate chicken DT40 cells lines with endogenously tagged AID, we first constructed vectors with a partial genomic sequence of AID with various 3' tags. For this, AID genomic sequence was placed into a pBluescript vector, followed by the construction of different tags: FLAG, Myc, FLAG-TEV-Myc and eGFP, and their ligation in frame with exon 5 at the 3' end of the AID coding sequence (Figure 28 and see 3.2.3).
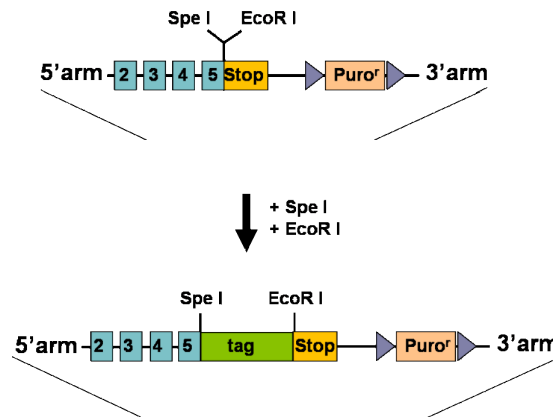
Figure 23. Schematic drawing of the construction of the tagged AID vectors. Different tags were ligated at the 3' end of a chicken AID gene in a Spe I and EcoR I double digested pBluescript plasmid. *loxP* sites, depicted as blue triangles, are located at the 5'- and 3'-end of a puromycin selection marker. 5'arm is 1 kb of the intron between exon one and exon two in the chicken gene and 3'arm stands for approximately 1 kb of the sequence downstream of the AID gene in the chicken genome. Exons are depicted as light blue, numbered squares. Tag marks either FLAG, Myc, FLAG-TEV-Myc or eGFP. Stop codon is depicted as a dark yellow square. Not drawn to scale.

FLAG and Myc tags were obtained by hybridizing two complementary oligos, followed by ligation into the expression vector containing AID genomic sequence. For generating FLAG-TEV-Myc, the tag was produced from 6 oligonucleotides with partially overlapping sequences at their ends to facilitate their ligation. eGFP was derived by PCR amplification from an eGFP expressing

plasmid (for details see 3.2.3.1). The vectors with inserts were verified both by digestion and by sequencing. The sequenced constructs were linearized with Not I before transfecting DT40 cells.

4.2.2.1.2 Transfecting DT40 for endogenous tagging of AID locus.

Since the initial transfection attempts with linearized constructs were not successful, indicated by the absence of puromycin resistant clones, the transfection procedure was modified. The DNA concentration used for the transfection was decreased from 50 µg to 25 µg, the final concentration of β-mercaptoethanol in the culture media was increased from 50 µM to 75 µM, and the cells were electroporated once instead of two short interval electroporations. This new transfection procedure generated 75, 70, 98 and 40 puromycin resistant clones for FLAG, Myc, FLAG-TEV-Myc and eGFP, respectively. Since some of the selection marker resistant clones were expected to have random integrations of the constructs into the genome, the clones were analyzed by southern blot. A southern blot probe was designed to anneal to exon 3 and intron 3 in the AID gene. The results from southern blot analysis of Spe I digested genomic DNA from puromycin resistant clones are shown in Figure 29.
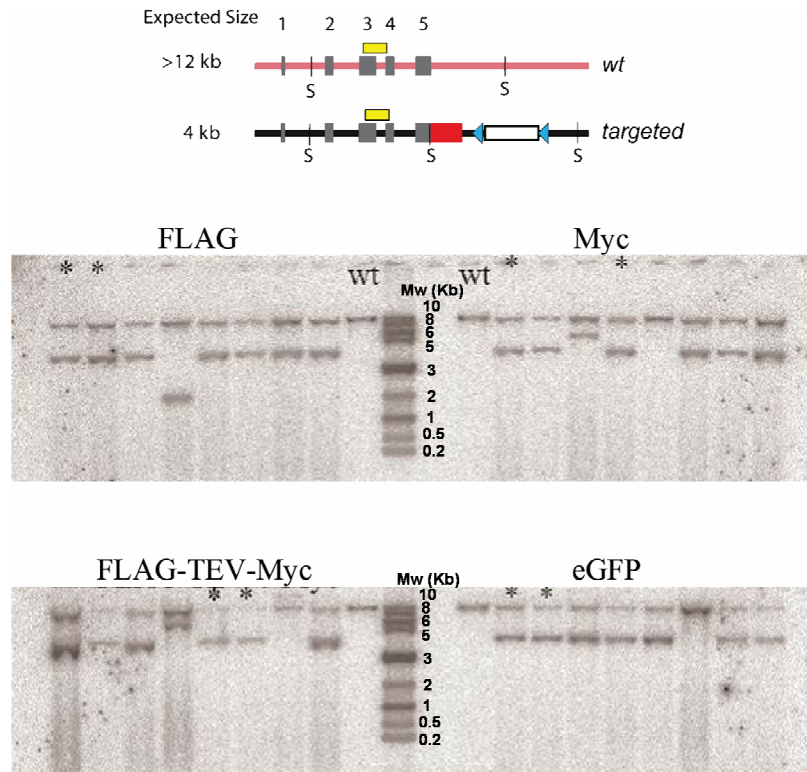
Figure 29. Southern blot analysis of puromycin resistant DT40 clones cut with Spe I. Schematic presentation of AID genomic locus. Grey boxes mark exons 1 to 5, red box depicts the tag, white box marks puromycin resistance cassette, yellow box marks the location of southern blot probe. Genomic DNA from puromycin resistant clones from transfections with AID-FLAG, -Myc, -FLAG-TEV-Myc and -eGFP were digested with Spe I and analyzed on a southern blot. Genomic DNA from wt cells is included as a control. Samples marked with '*' were later analyzed with a second southern blot.

The wild-type AID allele digested with Spe I yields a ~12 kb fragment, and the tagged AID allele yields a 4 kb fragment, as a Spe I restriction site lies immediately upstream of the tags. Therefore, samples with bands approximately 4 kb indicate targeted integration. Although a few samples reveal random integration, many of the samples contain a 4 kb band, e.g seven out of the eight eGFP and FLAG-tagged clones have a 4 kb band. Eight targeted integration (marked with '*' in Figure 29) were confirmed with a second southern blot. Here, the genomic DNA was digested with Mph1103I endonuclease, generating a 4.6 kb fragment for the wild-type allele and a 7.2-8.1 kb fragment for the tagged AID allele. The second southern blot verified that all of the chosen clones had a targeted integration (Figure 30).

121

Figure 24. Southern blot analysis of selected puromycin resistant DT40 clones with Mph1103I. Genomic DNA from selected puromycin resistant clones were digested with Mph1103I and analyzed on a southern blot. Genomic DNA from wt cells is included as a control.

4.2.2.1.3 Transient transfection of DT40 Cells with Cre Recombinase.

Eight clones confirmed to contain targeted integration of tagged AID, were transiently transfected (Amaxa electroporation) with a plasmid expressing Cre recombinase under the control of a CMV promoter. Expressed Cre recombines the two loxP sites, thereby removing the puromycin cassette. Cells were simultaneously transfected with a plasmid containing GFP under the same promoter and monitored for transfection efficiency after 30 hours. Despite a low transfection efficiency of the cells (< 10% GFP positive), clones were grown until sufficient density, and tested for puromycin sensitivity. The results from this test indicated a poor transfection efficiency, as we obtained 2/393 (puromycin sensitive clones/total number of clones) FLAG, 6/335 for Myc, 1/334 for eGFP, and no puromycin sensitive clones (0/323) for FLAG-TEV-Myc. The transfection method was therefore optimized, by using varying concentrations of plasmid DNA and a number of transfection programs as recommended by the manufacturer. The results of this optimization experiment indicated, that the program B023 and a plasmid amount of 3 µg should be used in the future for DT40 Cre transfection (other conditions indicated more than 2-fold lower transfection efficiencies). Southern blot analysis was carried out on the selection marker sensitive clones, to verify that the puromycin gene has been excised out of the genome. Two southern blot investigations of the puromycin sensitive

clones were carried out, using either Kpn I cleaved genomic DNA (Figure 31B) or Mfe I cleaved DNA (Figure 31C).

Figure 25. Southern blot analysis of puromycin sensitive clones. (A) Schematic representation of the AID locus in wt, targeted or Cre treated cells. Grey boxes indicated exons 1 to 5, red box marks the tag, white box marks puromycin cassette, yellow box depicts the location of the southern blot probe. Genomic DNA from puromycin sensitive clones was digested with either (B) Kpn I or (C) Mfe I and analyzed by a southern blot. Genomic DNA from wt cells is included as a control. The tag-controls are samples with digested genomic DNA from clones that still carry the puromycin gene. Sample in which the puromycin cassette was not excised, is indicated with "*".

Kpn I digestion of an AID allele containing a puromycin selection marker yields a 5.5 kb fragment, whereas an excised allele results in 3.1 kb. From Figure 31 we can deduce that 1/2 FLAG; 6/6 Myc; 1/1 GFP and 6/6 FLAG-TEV-Myc clones had excised the puromycin cassette. Mfe I digestion of an AID allele containing a puromycin selection marker yields a 4.4 kb fragment, whereas an excised allele results in 5.7 kb. Southern blot of Mfe I cleaved samples indicated that the selection marker had not been excised in one of the clones with a FLAG-tagged AID, confirming the findings of the Kpn I southern blot analysis. Therefore, 14 clones carrying tagged AID were derived - 6 with a Myc tag, 1 with an eGFP, 1 with a FLAG and 6 with a FLAG-TEV-Myc tag.

The eGFP signal from the clone carrying eGFP-tagged AID was compared to wild-type cells by flow cytometry, to confirm that the fusion protein is functional and expressed. The cells were stained with propidium iodide (PI), to exclude dead cells from the analysis (Figure 32). The FACS analysis confirmed the expression of eGFP from the endogenously tagged AID allele, due to an eight times higher than the background signal from wild-type cells.
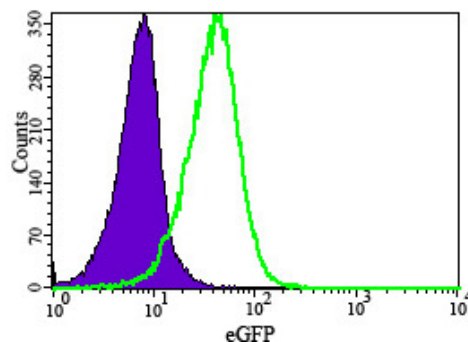
Figure 26. Analysis of eGFP-tagged AID DT40 by flow cytometry. Number of cells (y-axis) as a function of the GFP signal (x-axis, logarithmic scale). Results from wild-type cells are shown with a purple color whereas results from cells carrying eGFP-tagged AID are shown with a green color. Dead cells are excluded from the analysis by PI gating.

The FLAG, Myc and FLAG-TEV-Myc tagged cell lines were subjected to western blot analysis, to investigate whether the cells express the FLAG and Myc tagged AID. As an anti-Myc 9E11 antibody failed to detect the endogenously

tagged AID-Myc fusion protein, an anti-myc 9E10 antibody was used instead (Figure 33).
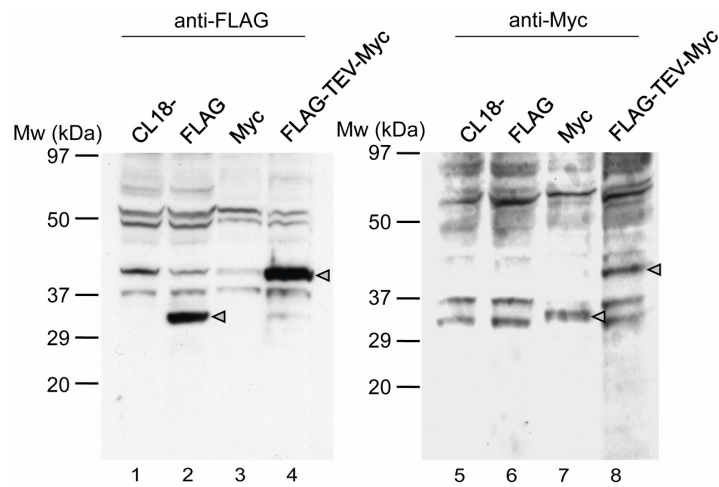
Figure 27. Western blot analysis of endogenously tagged AID in DT40 cells. (Left panel) DT40 cell lysates from control cell line (CL18-), FLAG, Myc or FLAG-TEV-Myc tagged cell lines, detected with anti-FLAG antibody. (Right panel) DT40 cell lysates from control cell line (CL18-), FLAG, Myc or FLAG-TEV-Myc tagged cell lines, detected with anti-Myc antibody. Triangles mark the presence of the tagged AID protein.

The results indicated that all three DT40 cell lines contained a functional tag and were expressed in these cells in comparable levels to each other. Furthermore, neither FLAG antibody nor Myc antibody cross-reacted with each other on AID. In conclusion, AID-tag DT40 cell lines were generated, that can be used for detecting endogenous AID protein and for investigating AID-interacting proteins.

4.2.2.1.4 Analysis of AID-eGFP localization in DT40.

In order to verify that the detected GFP signal was specific for AID-eGFP expression, we analysed the GFP signal under a fluorescence microscope. Wild-type cells and cells expressing AID tagged with eGFP were mixed in equal numbers and stained with the DNA specific fluorescent dye DRAQ5*TM* (Figure 34).

125

A) DRAQ5 **B) GFP** **C) Merge**
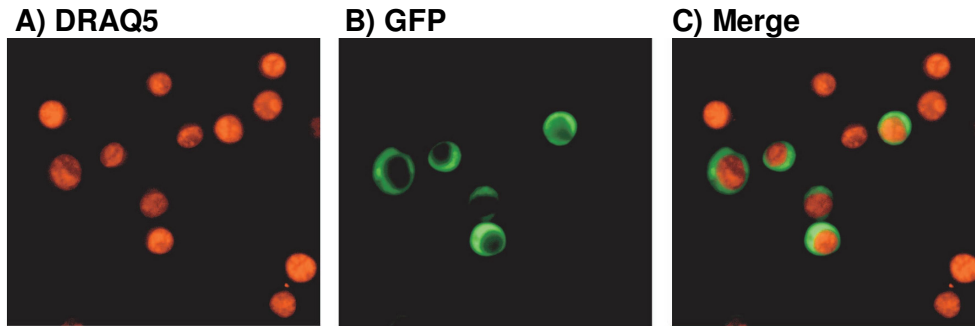
Figure 28. Endogenous expression of AID-eGFP fusion protein in DT40. Confocal images of eGFP and DRAG5*TM* signal from living wild-type cells mixed with cells expressing AID-eGFP. Equal numbers of wild-type and AID-eGFP DT40 cells were mixed and stained with DRAG5*TM*, followed by fluorescence microscopy.

Fluorescence microscopy confirmed that the detected GFP signal comes from the AID-eGFP allele, since a GFP signal was observed in approximately 50% of the cells (Figure 34). The cellular localisation of the fusion protein was investigated in the presence of leptomycin B (LMB), an inhibitor of CRM1-mediated nuclear export that has been reported to affect the cellular localization of overexpressed AID (Figure 35) (McBride et al., 2004).

A) Control **B) + LMB**
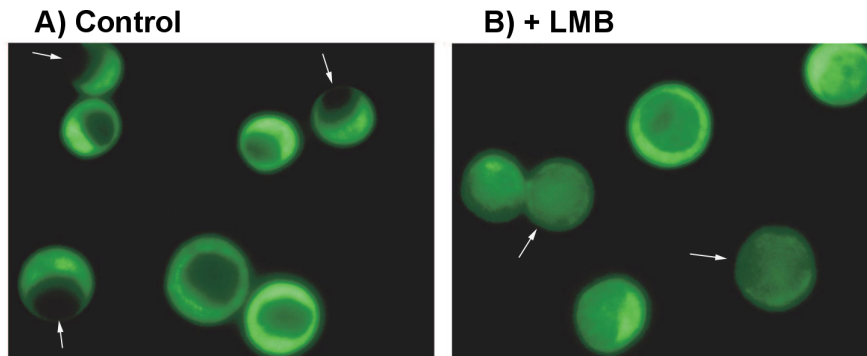
Figure 29. AID localization in AID-eGFP DT40 cells after Leptomycin B treatment. Live AID-eGFP DT40 cells were analyzed by fluorescence microscopy in the A) absence or B) presence of LMB. Cells were treated with 25 ng/mL LMB for 5 hours. White arrows indicate cells that have a significant redistribution of the eGFP signal.

Endogenously expressed C-terminally tagged AID resides predominantly in the cytoplasm as indicated by the eGFP signal (Figure 35A). After treating the cells with LMB, the eGFP signal partially relocates into the nucleus (Figure 35B). Even though the AID-eGFP fusion protein remains to be confirmed by western blotting, our results already suggest that endogenously tagged AID enters the nucleus and is actively exported from the nucleus by a CRM1-dependent nuclear export pathway. However, the translocation of the signal into the nucleus was not complete, because a significant proportion of the GFP signal remained in the cytoplasm. Therefore, the eGFP signal

4.2.2.2 Identification of AID-interacting proteins.

DT40 cells that contain AID tagged with FLAG, Myc or FLAG-TEV-Myc, can be used for biochemical studies of AID and for identifying its interaction partners. As the knowledge of proteins that interact with AID in a physiological environment is scarce, we aimed to use the cells in unbiased coimmunoprecipitation studies for identifying factors that form a complex with AID.

4.2.2.2.1 Optimization of AID immunoprecipitation conditions in DT40 cells.

In order to use our tagged AID cell lines for identifying AID-interacting proteins, it was necessary to optimize the experimental conditions. This comprised of various parameters, including cell lysis, number of cells per volume lysate, the volume of beads per volume lysate and washing conditions.

First, cell lysis was tested using either 0.3% Triton X-100 or 0.3% NP40 as a detergent in the lysis buffer (Figure 36). The cell lysates from DT40 AID-FLAG-TEV-Myc were subjected to immunoprecipitation (see 3.2.4.3), using commercial anti-FLAG antibody crosslinked to agarose. Following the binding step, the beads were washed and the protein was eluted with a 3xFLAG peptide using a concentration recommended by the manufacturer. The comparison of AID-tag protein signal from total cell lysates (Figure 36 lane 1-4) and from the flow-through (Figure 36 lane 5-6) of the pull-down indicated that AID-tag fusion protein was efficiently bound to the anti-FLAG affinity beads. The protein

elution fractions showed that less than 50% of the protein remained bound to the beads (Figure 36 lane 7-8 vs 9-10). There was also some contamination of the Ig heavy and light chains from anti-FLAG affinity beads, seen as a slower and faster migrating band compared to AID-tag protein (Figure 36 lane 7), indicating that bead removal conditions needed optimization. Nevertheless, the results indicated that both lysis conditions worked with a similar efficiency.
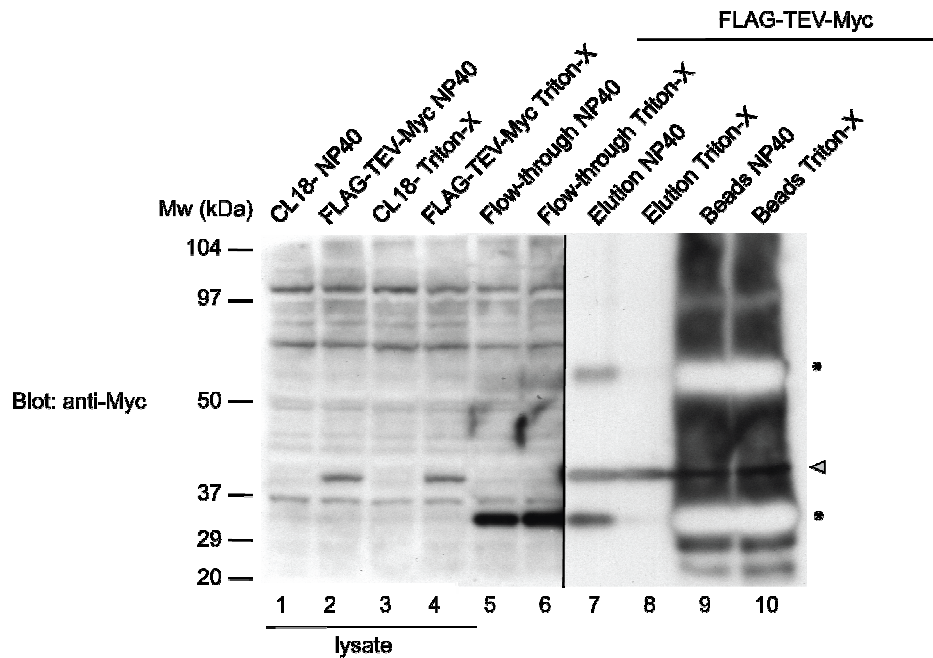
Figure 30. Optimization of DT40 lysis conditions for AID pull-down. Cells were lysed with lysis buffer containing either 0.3% NP40 or 0.3% Triton-X, subjected to immunoprecipitation with anti-FLAG affinity beads and detected with anti-Myc antibody. The corresponding fractions of the pull-down experiment were compared to each other in order to estimate the efficiency of each step in the experiment. !/100 of total was loaded on lane (1-6), 1/4 of total was loaded (7-10). The triangle marks AID-FLAG-TEV-Myc fusion protein. The asterisks mark Ig heavy and light chain contaminants from the anti-FLAG affinity beads. The film was exposed shorter for lanes 7-10 compared to lanes 1-6, since the signal was too strong at longer time of exposure.

Even though, silver stain analysis of the eluates from the pull-downs confirmed that both lysis conditions were compatible with AID-FLAG-TEV-Myc immunoprecipitation, we chose to proceed with 0.3% Triton X-100 in future experiments, because Triton X-100 can be used in broader ranger of concentrations without affecting the functionality of anti-FLAG beads.

Next, we optimized the bead volume per cell number. By using decreasing volumes of anti-FLAG beads during the pull-downs, we determined that 2.5 μl of anti-FLAG affinity beads was optimal per 5 x 10^7 cells (data not shown), thus allowing for the scaling of the experiment while keeping the bead to cell number ratio. In order to use the Myc epitope for immunoprecipitation of AID in AID-Myc and AID-FLAG-TEV-Myc cell lines, anti-Myc affinity gel had to be generated. The 9E10 anti-Myc antibody was cross-linked to cyanogen bromide activated sepharose, and the remaining active groups were then blocked. To test bead volume to cell number, various volumes of anti-Myc affinity beads were used for AID pull-down from AID-FLAG-TEV-Myc DT40 cell line (data not shown). Based on this experiment, we decided to use 15 μl of anti-myc beads per 5 x 10^7 cells. The relative difference of the optimal volumes for anti-FLAG and anti-Myc affinity beads might be due to both, the lower affinity of anti-Myc antibody towards its epitope, as well as small differences in the cross-linking efficiency of the antibodies to the beads. For determining the efficiency of anti-Myc affinity beads, we compared the protein signal intensities for total cell lysates, flow-through after binding to the beads, the percentage of protein eluted from the beads and the percentage which remained bound to the beads after elution, on a western blot (Figure 37). In parallel, we tested 1xMyc (Figure 37 lanes 2, 5, 7, 9) and 3xMyc peptides (Figure 37 lanes 4, 6, 8, 10) for their ability to elute (Figure 37 lane 7 vs 8) our protein.
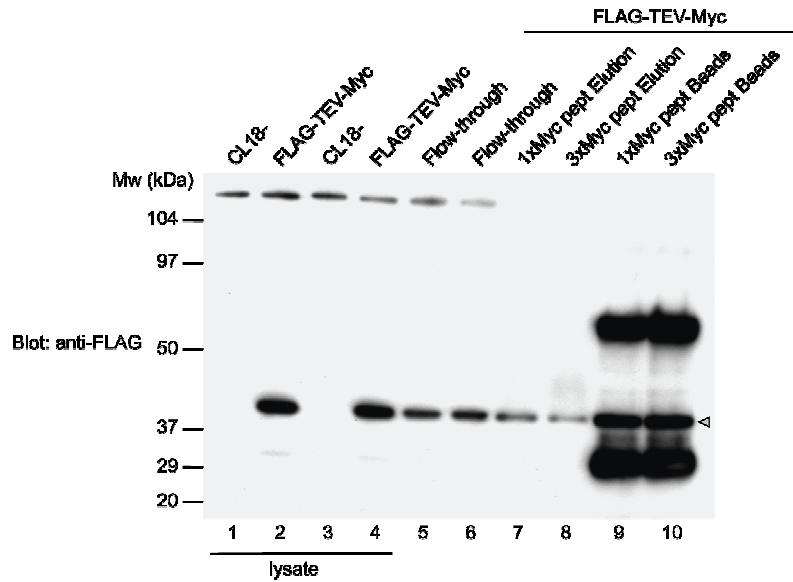
Figure 31. Optimization of AID pull-down from DT40 with anti-Myc affinity beads. Protein pull-downs were carried out from DT40 extracts using anti-myc affinity beads. Pull-down was carried out using either a 1xMyc peptide (odd lane numbers) or 3xMyc peptide (even lane numbers) for protein elution from the beads. The efficiency of these pull-downs was estimated with a western blot using anti-FLAG antibody for detecting the protein. CL18- DT40 cell line was used as a control. The triangle marks AID-FLAG-TEV-Myc fusion protein.

Comparing anti-FLAG and anti-Myc affinity gels to each other (Figure 36 vs Figure 37) led us to conclude: 85-90% of the total AID-FLAG-TEV-Myc pool was bound to the anti-FLAG beads and approximately 50-60% of the bound molecules were eluted by the 3xFLAG peptide. On the other hand, the anti-Myc affinity gel indicated that 40-50% of the total AID-FLAG-TEV-Myc pool was bound to the beads and approximately 20% of the bound molecules were eluted by the Myc peptide. Thus, the elution efficiency was not as high as for the 3xFLAG peptide, and did not improve when 3xMyc peptide was used for elution.

Having established the optimal bead volumes for both affinity gels and elution conditions, we continued with improving other pull-down parameters such as the signal to noise ratio. To decrease the background binding of unspecific proteins to the beads, higher salt concentrations, addition of 1xFLAG peptide in the washes, pre-clearing and post-clearing with anti-Ig affinity gel were tested (data not shown). The results indicated that the addition of 1xFLAG peptide in the washes improves the results, as it removed non-specific sticking of proteins to the beads, at the same time not eluting AID-FLAG from the beads. On

the other hand, pre-clearing of the samples with an anti-Ig affinity gel had only minor effects on the non-specific binding. This might be due to a strong excess of proteins in the lysate, which are not removed by pre-clearing and therefore stick to the anti-FLAG beads. Post-clearing of the eluates (data not shown) indicated that the signal for AID-FLAG fusion decreased together with non-specific proteins. Therefore, both pre-clearing and post-clearing in general, did not significantly improve the signal to noise ratio during the pull-down experiments. By increasing the salt concentration in the bead washes, it was possible to remove background sticking, but it also led a loss in AID-tag binding to the beads. Furthermore, as high salt concentrations might disrupt the binding of specific interaction partners to AID, we chose to continue using 150 mM NaCl in the washes.

In order to use the AID-FLAG-TEV-Myc cell line for two-step pull-downs (first IP with anti-Myc, followed by anti-FLAG IP), it was necessary to test whether the TEV cleavage sites in the double tag are functional. Therefore, pull-downs were carried out, using anti-Myc affinity gels to bind the protein and then cleave with various amounts of TEV protease (Figure 38).
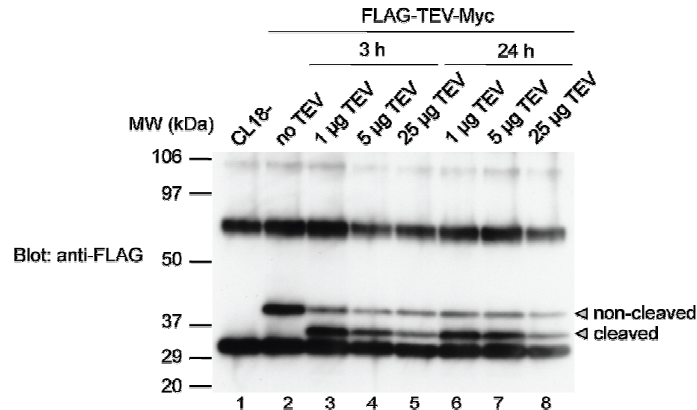
Figure 32. Optimization of TEV cleavage. Tagged DT40 cells were subjected to protein pull-downs using anti-Myc affinity beads and analyzed by western blot. The resin bound AID-FLAG-TEV-Myc proteins were treated with various concentrations of TEV protease at 3 h and 24 h time-points. Immunoprecipitated proteins were detected with anti-FLAG antibody in order to visualize both the non-cleaved and cleaved (marked with arrows) fractions of AID on the beads. The CL18- DT40 cell line was used as a control.

Incubating with TEV protease resulted in the appearance of an additional lower-migrating band and the simultaneous disappearance of the higher-migrating non-cleaved band. The efficiency of the cleavage was already maximal at 1 μg TEV, and the subsequent experiments revealed that a 0.1 μg of TEV per 5 x 10^7 cells for 3 h is suitable for efficient elution of the tagged AID protein.

In conclusion, by optimizing the immunoprecipitation conditions, we determined suitable parameters for identifying AID-interacting proteins in large-scale experiments.

4.2.2.2.2 Mass spectrometry.

By analyzing the disribution of AID protein in the cytoplasm, nucleoplasm and chromatin, we found that the majority of the total AID pool is localized in the cytoplasm (approximately 80-90%), 5-10% nucleoplasm and 5-10% is in the chromatin fraction (Figure 39). This is in agreement with our observations in DT40 AID-eGFP cell line as well as earlier studies which also found AID predominantly in the cytoplasm (Aoufouchi et al., 2008; McBride et al., 2004; Rada et al., 2002).
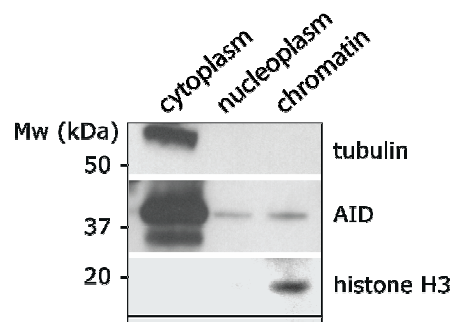
Figure 33. Sub-cellular localization of AID in DT40. Cytoplasmic, nuclear and chromatin fractions (see 3.2.4.2) were isolated from DT40 AID-FLAG-TEV-Myc cells and analysed for relative AID signal by using anti-FLAG antibodies. Lanes contain equivalent loadings. Tubulin and histone H3 were included as internal controls to assess the purity of the corresponding fractions.

In order to identify proteins in complex with AID, cells were grown to large volumes (20L), followed by the isolation from cytoplasmic fraction, nucleoplasm and chromatin. AID was immunoprecipitated from these sub-cellular fractions and the proteins binding to AID were identified by mass spectrometry (Figure 40

and 41). In parallel, samples were treated with RNAse prior to IP to degrade RNA and thus abolish interactions mediated by RNA. Proteins identified from the CL18- control IP were considered as nonspecific and disregarded.
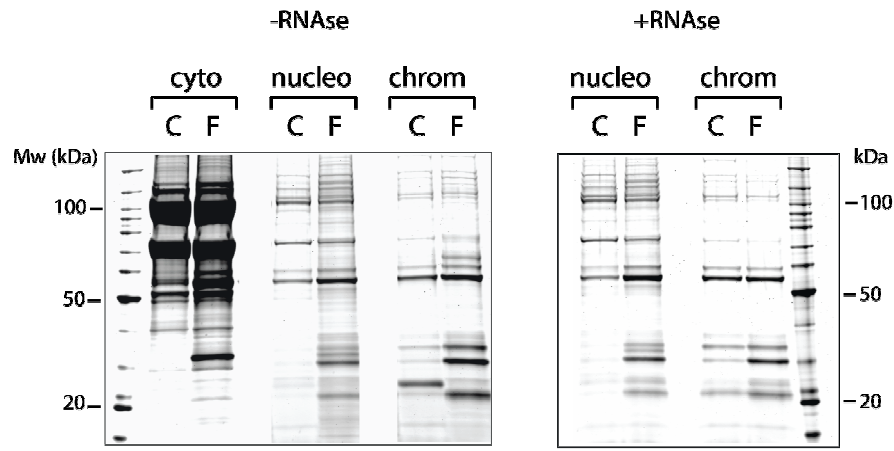
Figure 34. Identification of AID-interacting proteins. Cytoplasmic, nuclear and chromatin fractions were isolated (see 3.2.4.2) from DT40 AID-FLAG-TEV-Myc cells, immunoprecipitated by anti-FLAG beads (see 3.2.4.3) and visualized on a 4-12% gradient gel by Sypro Ruby staining (see 3.2.4.5). Left panel indicates samples not treated with RNAse A, whereas right panel shows samples treated with RNAse A prior to immunoprecipitation. C marks CL18- cells that were used as a control; F marks AID-FLAG-TEV-Myc cells.

Of the proteins identified, I will discuss briefly some interesting, albeit unverified proteins (Figure 41 and Appendix Table 1). Proteins identified in the cytoplasm in the absence of RNAse treatment indicated various factors involved in translation, including several ribosomal proteins. We also found factors involved in nucleo-cytoplasmic transport, such as importin subunit beta-1 and GTP binding protein Ran, suggesting that these proteins might be involved in the transport of AID into the nucleus. In addition, exportin 5 was identified as an AID-binding candidate, suggesting its involvement in the export of AID from the nucleus. However, its interaction with AID requires confirmation in future studies. Mass spectrometry also revealed several proteins known as components of various signaling pathways such as RACK1, which binds activated Protein kinase C (PKC) and the substrates of PKC, thereby acting as an adaptor for PKC mediated phosphorylation; or 14-3-3 protein, involved in a number of signaling pathways.

In the nucleoplasm (in RNAse treated sample), we identified several signaling proteins such as nuclear calmodulin-binding protein, interleuking

enhancer binding protein; a few DNA repair proteins such as RAD54-like 2 and Mre11, however these interaction candidates gave a relatively low number of identified peptides and thus a weak signal. As for the cytoplasmic fraction, a number of RNA binding proteins were identified, such as splicing factor 3B subunit 1, 3 and 4, pre-mRNA-processing factor 6, pre-mRNA-processing-splicing factor 8 and DEK oncogene that binds both DNA and RNA.

In chromatin fraction (RNAse treated sample), we found a significant number of proteins known to be part of the transcription elongation machinery. These include RNA polymerase II-associated factor 1, transcription elongation factors SPT5 and SPT6, FACT complex subunits and splicing factor 3B subunits 1, 3 and 4. In addition, mass spectrometry also revealed some DNA repair factors such as Rad16 and Topoisomerase I.

Cytoplasm		Nucleoplasm		Chromatin	
RACK1	-	Interleukin enhancer binding protein	+/-	Highly divergen homeobox	-
14-3-3 protein beta/alpha	-	Nuclear calmodulin-binding protein	+/-	Interleukin enhancer binding protein	+/-
Tumor supressor p53	-	Nuclease-sensitive element-binding prot 1	+/-	Nuclear calmodulin-binding protein	+/-
Survival motor neuron protein	-	Cytoplasmic activation-proliferation-associated protein	-	Nuclease-sensitive element-binding prot 1	-
Interleukin enhancer binding protein	-	Matrin 3	+/-	RNA polymerase II-associated factor 1	+/-
Poly(rC)-binding protein 2	-			Transcription elongation factor SPT5	+
Pre-mRNA-processing splicing fact 8	-	DEK oncogene (DNA binding)	+/-	Transcription elongation factor SPT6	+
Poly(A) binding protein	-	Splicing factor 3B subunit 1	+/-	pre-mRNA-processing factor 6	+/-
Ribosomal proteins	-	Splicing factor 3B subunit 3	+/-	FACT complex subunit SPT16	+
		Splicing factor 3B subunit 4	+/-	FACT complex subunit SSRP1	+
Importin subunit beta-1	-	Pre-mRNA-processing factor 6	-	ATP-dependent RNA helicase DDX1	-
GTP-binding protein Ran	-	Pre-mRNA-processing-splicing factor 8	+/-	DEK oncogene (DNA binding)	+/-
		Nuclear cap-binding protein subunit 1	+	Nucleolin	-
Exportin 5	-	Nucleolin	-	Poly(A) binding protein	-
		Poly(A) binding protein	-	Splicing factor 3B subunit 1	+/-
		ATP-dependent RNA helicase DDX1	-	Splicing factor 3B subunit 3	+/-
				Splicing factor 3B subunit 4	+/-
		RAD54-liko 2	-		
		Endo/exonuclease Mre11	-	Poly(ADP)-ribose polymerase 1	-
				Topoisomerase I	+/-
				DNA excision repair protein Rad16	-

Legend

■ Signaling pathways

■ RNA binding and processing, splicing

■ DNA repair

■ Import into nucleus

■ Export from nucleus

Figure 35. AID-interacting protein candidates identified by mass spectrometry. The listed proteins were identified from cytoplasm (-RNAse), nucleoplasm (+/-RNAse) or chromatin (+/-RNAse). The colour of the letters represents the known function of these candidates as indicated in the legend. + or – behind each candidate indicates, whether it was identified in RNAse treated sample or not treated sample, respectively. Candidates found in both samples are marked as +/-.

By comparing the proteins from all three cellular fractions, we found significant differences between the cytoplasm, nucleoplasm and chromatin

(Figure 41 and 42). On the other hand, nucleoplasm and chromatin fractions showed more resemblance to each other than to the cytoplasm. These results form the basis for further studies that address the question, which of the identified factors are direct binding partners for AID and how they are involved in the process of AID-mediated DNA deamination.
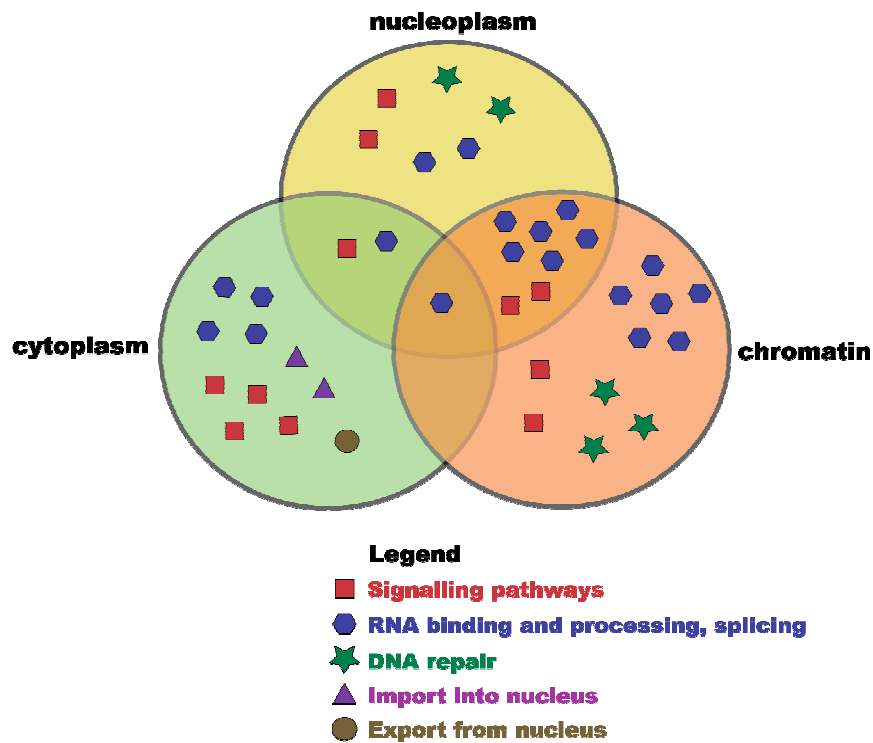
Figure 36. The intracellular distribution of AID-interacting candidates. Cytoplasm, nucleoplasm and chromatin are depicted as coloured circles. AID-interacting candidates identified by mass spectrometry are marked as in the figure legend and their place in the circle depends on the cellular fractions where these candidates were identified.

4.2.2.2.3 Confirmation of AID-interaction candidates.

In order to confirm the interaction of AID with some of the candidates, we performed cell fractionations, followed by immunoprecipitation of AID and subsequent detection of the interaction candidate by western blot (Figure 43). Western blot indicated the immunoprecipitation of Poly(ADP)-ribose polymerase (PARP) in the chromatin fraction (Figure 43A), however this interaction seemed to be RNAse sensitive since the signal was lost in the sample that was treated

with RNAse A prior to immunoprecipitation. Using an antibody against Poly(ADP)-ribose chains (Figure 43B), we observed several bands in the cytoplasmic fraction of the AID pull-down and a distinct band in the chromatin fraction, which was moving at the same molecular weight as PARP and was absent in RNAse A treated sample. This suggested that at least a proportion of PARP molecules in a complex with AID is Poly(ADP)-ribosylated.

In contrast, the signal of RACK1 indicated an interaction with AID in the cytoplasm and this seemed to be resistant to RNAse A treatment, suggesting that this might be a direct protein-protein interaction (Figure 43C). Using this approach, other candidates identified in a complex with AID will be investigated, determining interactions that are mediated via RNA or proteins.
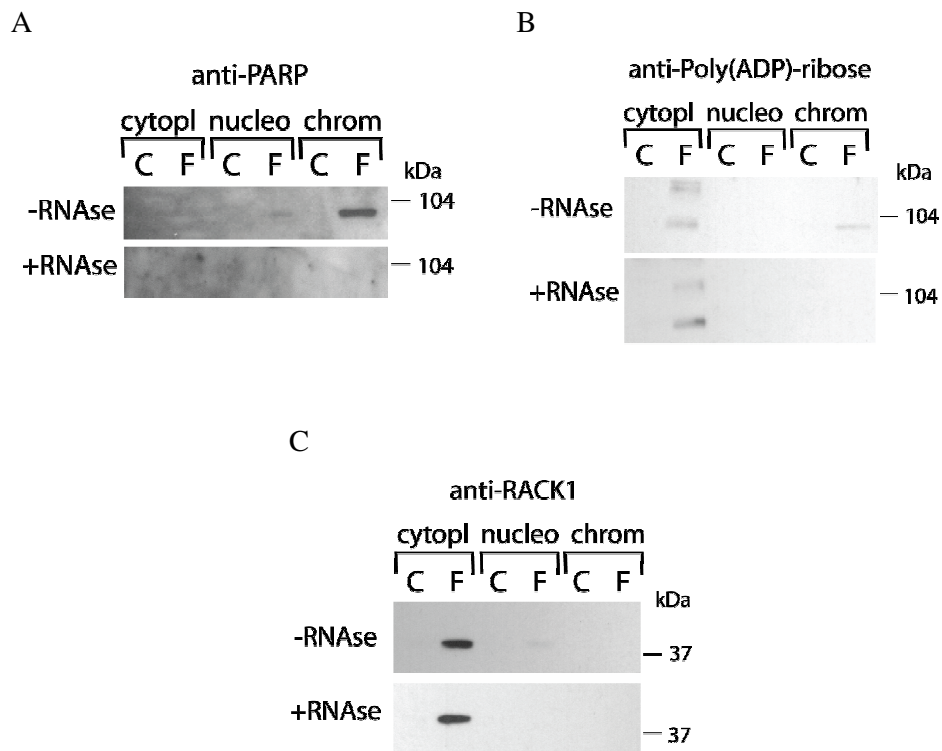
Figure 37. Confirmation of AID-interacting proteins. AID-FLAG-TEV-Myc and CL18- DT40 cells were subjected to the isolation of cytoplasm, nucleoplasm and chromatin. The samples were either threated with RNAse A or not, followed by immunoprecipitation of AID with anti-FLAG affinity beads. The eluates were analysed by western blot, using either anti-Poly(ADP)-ribose polymerase antibodies (A), anti-Poly(ADP)-ribose antibodies (B) or anti-RACK1 antibodies (C). C marks CL18- as a control pull-down for each fraction, F marks AID-FLAG-TEV-Myc pull-down.

4.2.3 Discussion.

By constructing four vectors with partial AID locus and a tag sequence, followed by introducing the vector in the endogenous AID locus via homologous recombination, we generated four DT40 cell lines – an AID-eGFP, AID-3FLAG, AID-3Myc and AID-3FLAG-2TEV-3Myc. These cell lines were confirmed by southern blot analysis (Figure 31), FACS analysis (Figure 32) and western blot (Figure 33). Thus, new DT40 cell lines were obtained in which one AID allele is tagged, leaving the fusion protein under the regulation of its endogenous promoter.

Fluorescence microscopy showed that AID with an eGFP tag at the C-terminus localizes predominantly in the cytoplasm (Figure 34). A similar AID distribution was found in AID-FLAG-TEV-Myc cells (Figure 39). These findings partially confirm earlier observations (Ito et al., 2004), which analyzed overexpressed C-terminally GFP tagged AID in mouse fibroblast, as well as the recent (McBride et al., 2004) investigation of the sub-cellular distribution of overexpressed C- and N-terminally GFP tagged AID in mouse B-cells. Both studies also reported, that treating the cells with LMB, a specific inhibitor of CRM1, results in nearly complete nuclear accumulation of the GFP-tagged AID. However, LMB treatment of our cell line indicated a more modest retention of AID in the nucleus, since the GFP signal intensity was still higher in the cytoplasm than in the nucleus (Figure 35). Furthermore, the cells did not respond in a homogeneous manner to LMB treatment, suggesting a complex regulation of AID localization. This discrepancy is likely to be a result of analyzing overexpressed versus endogenously expressed AID, and underlines the importance of our DT40 cell lines that possess endogenously expressed AID, representing a more physiological condition compared to overexpression.

The fact that most of AID is located in the cytoplasm and a small fraction is present in the nucleoplasm and on chromatin, makes it difficult to identify AID-interacting proteins in the nucleoplasm and chromatin fractions. However, this was overcome in our study by using large volumes of cells (e.g. 20 Litres). Using co-immunoprecipitation and subsequent mass-spectrometry, we have identified proteins that form a complex with AID in the cytoplasm, nucleoplasm and on chromatin (Figure 41 and Appendix Table 1). Proteins found to be in a complex

with AID include factors that participate in nucleo-cytoplasmic transport, intracellular signaling and DNA repair. In the cytoplasm, a number of factors (components of translation machinery and RNA-binding factors, including ribosomal proteins) overlap with proteins identified in a complex with APOBEC3G (Kozak et al., 2006), a member of the AID/APOBEC family of proteins. The identification of these mRNA binding proteins and components of the translation machinery is also indirectly supported by a recent finding, suggesting that AID can bind to mRNA (Nonaka et al., 2009), as reported earlier for APOBEC3G (Kozak et al., 2006).

Strikingly, in the chromatin fraction, AID resides in a complex containing a significant number of RNA-binding proteins involved in transcription elongation and splicing (Figure 41). This suggests, that transcription elongation and/or RNA processing might be important for the targeting and execution of DNA deamination. Although, the participation of transcription in AID-induced deamination has been suggested earlier (Nambu et al., 2003), an unbiased approach using endogenously expressed AID for identifying its binding partners, as our DT40 system, has been missing so far. Interestingly, the mutation pattern in AID-targeted Ig loci starts approximately 150 bp downstream of transcription initiation site, is highest after this region and gradually decreases over the next 1.5 kilobase (Both et al., 1990; Lebecque and Gearhart, 1990; Longerich et al., 2006; Rada and Milstein, 2001). This resembles the general occupancy of RNA polymerase II on a transcribed locus, since a proportion of RNA polymerase II is known to fall off the chromatin during transcription, mainly due to nucleosomes that interfere with the movement of elongating RNA polymerase (Kristjuhan and Svejstrup, 2004).

Using Ingenuity Pathway Analysis software and analyzing proteins identified specifically in the chromatin fraction of AID IP (RNAse treated), we were able to build a protein interaction network, based on protein interaction databases (Figure 44). Many of the factors identified in our study, are known to bind each other, suggesting the presence of a large complex. In the future, it will be important to analyse AID-interacting candidates further – to examine which residues in AID mediate a particular protein-protein interaction. This would be particularly interesting with the factors found on chromatin, since AID-binding candidates identified in this fraction suggest a complex involved in transcription

elongation and RNA processing and might give a mechanistic view of the process of AID-mediated DNA deamination.
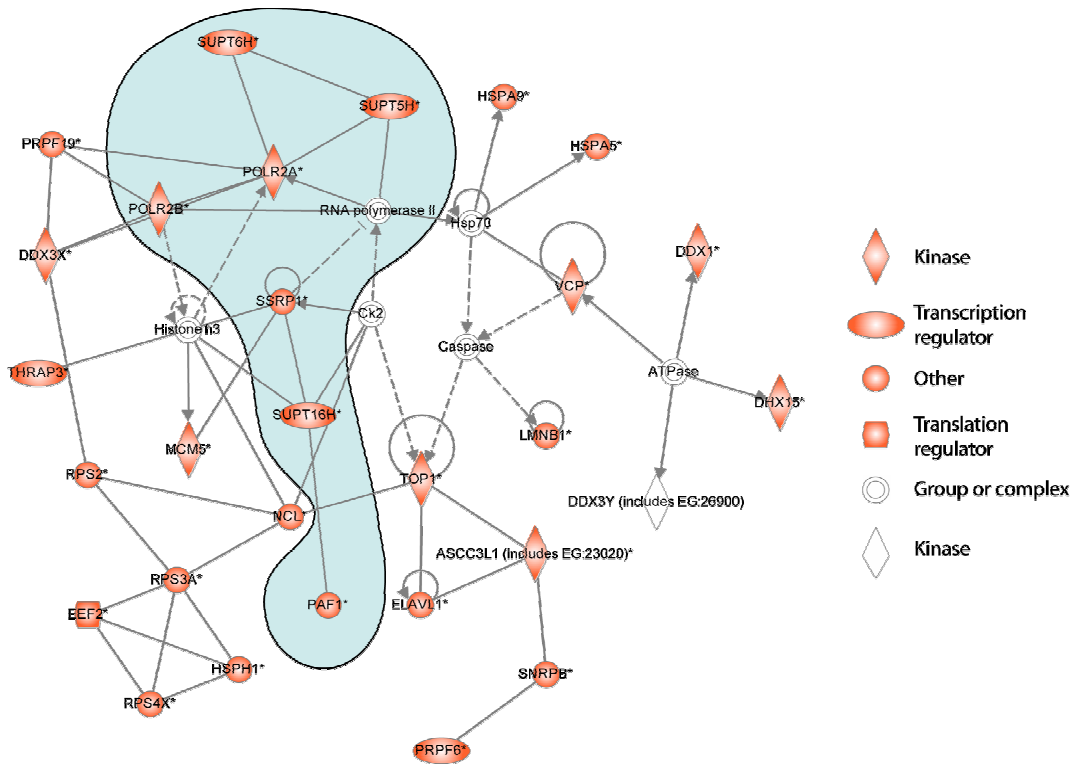
Figure 38. Protein-protein interaction network for AID-interacting candidates in the RNAse treated chromatin fraction. Proteins identified by mass spectrometry (see also Appendix Table 1) were searched for known interactions using information from protein interaction databases. Each line depicts a previously identified interaction between the candidates. The proteins in red mark the candidates identified in our mass spectrometry analysis. White marks proteins not found in our samples but previously identified to interact with red candidates. Light blue highlights components of elongating RNA polymerase.

Recently, it has been found that the V-region and switch regions, both targeted by AID, are being transcribed in sense and anti-sense direction (Perlot et al., 2008; Ronai et al., 2007). It is possible that this might help to explain the lack of strand biases in AID-mediated mutations since transcription bubble would make the ssDNA accessible for DNA deamination (Figure 45). Future experiments will hopefully clarify how transcription of a locus affects AID-mediated DNA deamination. In addition, AID has been shown to directly interact with CTNNBL1 (Conticello et al., 2008), a protein known to bind other splicing

factors. Since my study found a number of splicing factors in a complex with AID in the nucleus, it is possible that RNA processing might also play a role in the DNA deamination process.

A)

B)
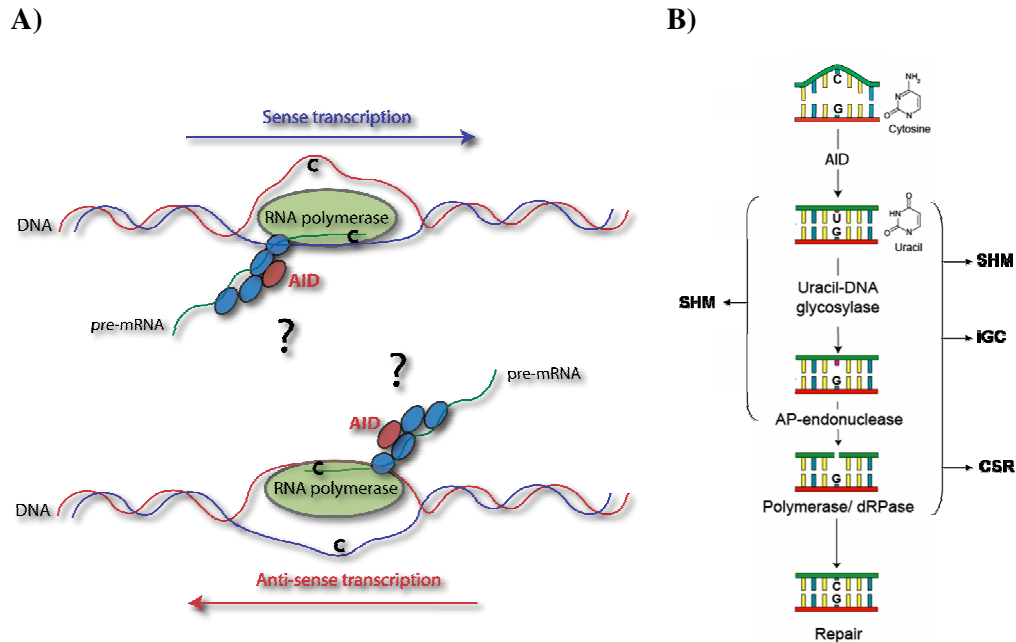
Figure 39. A model for AID-mediated DNA deamination. A) A schematic depiction of AID-mediated DNA deamination, where AID binds to elongating RNA polymerase II during both sense and anti-sense transcription in the Ig locus. This opens the chromatin and makes ssDNA accessible for cytosine deamination by AID. AID is depicted in red. B) Schematic depiction of AID-mediated DNA deamination pathway. AID induced cytosine deamination results in uracil, which can initiate SHM, iGC or CSR.

In conclusion, the unbiased identification of AID-interacting proteins by mass spectrometry forms the basis for future analysis of the identified factors and the investigation of their role in the process of AID-mediated DNA deamination. It also forms the basis of a mechanistic view of AID-mediated deamination and opens up new possibilities to tackle the process of enzymatic DNA deamination.

4.3 Dissecting downstream DNA repair pathways of AID-mediated deamination.

4.3.1 Introduction.

4.3.1.1 DNA repair.

The repair of DNA damage is essential for the survival of an organism. Genotoxic factors can stem from exogenous as well as endogenous sources. These include chemical agents (oxygen, radicals, or alkylating molecules), physical agents (e.g. UV- or gamma-radiation) and enzymes such as AID. DNA damage itself comes in different forms and includes base damage, inter and intra-strand DNA cross-links and single- and double-strand DNA breaks [Reviewed in (Lindahl, 1993)]. Depending on the cellular context as well as the location of the lesion, AID-induced uracils can initiate various DNA repair pathways, however, the choice of which DNA repair pathway is used, is still poorly understood. If AID deaminates cytosines in switch regions in activated B-cells, it induces class switch recombination via NHEJ. On the other hand, if AID targets the variable region, U lesions can be processed either by BER or MMR, followed by repair via an error free or error prone manner (Neuberger et al., 2003; Odegard and Schatz, 2006; Rada et al., 2004; Rada et al., 2002). In DT40, AID induces lesions in V-region that are preferentially repaired via gene conversion (Arakawa et al., 2002; Harris et al., 2002b). These differences in downstream processing of AID lesions are at least partially dependent on the genomic loci, since some genes targeted by AID readily accumulate mutations whereas other AID-targeted loci undergo error-free DNA repair (Liu et al., 2008). This also suggests that the repair of deaminated cytosines generated by spontaneous DNA deamination, can have different outcomes depending on the cellular context and genomic location of the lesion.

We therefore aimed to study the fate of dU lesions in meiotic cells, by using AID as a source for uracils in DNA. We chose to utilize cells undergoing meiosis in two model organisms, the fission yeast (*S. pombe*) and the nematode (*C. elegans*) (described below). Both organisms do not contain endogenous DNA

deaminases (Conticello et al., 2005), allowing us to detect changes caused by AID expression. We aimed to express AID during meiosis in a Spo-11 deficient background (Rec12 in *S. pombe*), which is devoid of meiotic DSBs (described below in more detail). Since Spo-11 deficiency leads to a lack of meiotic recombination (De Veaux et al., 1992; Dernburg et al., 1998), chromosome segregation defects, and loss in viability in both organisms, it is possible to detect effects of AID-induced U:G mismatches on meiosis and to dissect the DNA repair pathways involved in the process.

4.3.1.2 Base excision repair.

Base excision repair mechanism recognizes, excises, and accurately replaces specific forms of base modifications [Reviewed in (Barnes and Lindahl, 2004)]. The first step of BER typically involves the recognition and excision of an inappropriate base, such as uracil, 3-methyladenine, 8-OH-dG, or formamidopyrimidines, with each particular lesion being recognized by a specific DNA glycosylase (Lindahl and Wood, 1999). They catalyze hydrolysis of the *N*-glycosylic bond, resulting in an abasic site with an intact DNA phosphodiester backbone. Next, apurinic endonuclease 1 (APE1) both incises immediately 5′ to an AP site and removes 3′-obstructive termini (Demple and Sung, 2005; Wilson and Barsky, 2001). DNA polymerase β (POLβ) replaces the excised damaged nucleotide and removes the 5′-terminal abasic fragment left behind by APE1 incision (Bennett et al., 1997; Mol et al., 2000). Once the DNA terminal ends have been processed to contain a 3′-hydroxyl group and a 5′-phosphate moiety, the remaining nick is sealed by either a complex of X-ray cross-complementing 1 (XRCC1) protein and DNA ligase 3α (LIG3α), or the individual protein DNA ligase 1 (LIG1) (McKinnon and Caldecott, 2007). Even though it is well established that BER is important in somatic cells, including Ig diversification in B-cells, its role in fixing base lesions in meiotic cells is less well studied.

4.3.1.3 Meiosis.

Meiosis is a cell division process in sexually reproducing organisms that results in the generation of gametes. It is characterized by two consecutive cell divisions that give rise to four haploid cells from a diploid zygote. After meiotic recombination and during the first reductional division, paired homologous chromosomes are separated. The second meiotic division is similar to mitosis, where sister chromatids are passed to daughter cells, ultimately resulting in four cells that are genetically different from each other. Meiotic recombination plays an important role in the establishment of genetic differences of daughter cells [Reviewed in (Page and Hawley, 2003)] – it shuffles alleles between homologous chromosomes, exerting a profound effect on genetic diversity. At the same time, meiotic recombination provides physical contacts (chiasmata) between homologous chromosomes, thereby facilitating the proper segregation of chromosomes at the first meiotic division [Reviewed in (Kauppi et al., 2004)].

Meiotic recombination is initiated by the formation DNA double-strand breaks (DSBs) and subsequent utilization of DSB repair. DSBs are generated directly through a topoisomerase-like transesterase reaction by Spo-11, a meiotically induced protein that has sequence similarity to an archeal type II topoisomerase (Keeney et al., 1997; Klapholz et al., 1985; Lin and Smith, 1994). Meiotic recombination also involves eukaryotic *Rad51* – a functional homolog of the *Escherichia coli* strand transfer enzyme RecA (Shinohara et al., 1997; Shinohara et al., 1992). Spo-11 mutants are unable to accomplish meiotic recombination and exhibit defects in chromosome segregation, leading to aneuploid gametes, causing infertility in mammals, and a decrease in viability of spores in fission yeast (Keeney et al., 1997; Klapholz et al., 1985; Lin and Smith, 1994; Whitby, 2005). Because Spo-11 homologs are widely conserved in eukaryotes and are essential for meiotic recombination, DSB formation is most likely a universal mechanism for initiating meiotic recombination (Baudat et al., 2000; Romanienko and Camerini-Otero, 2000).

4.3.1.4 Schizosaccharomyces pombe and Caenorhabditis elegans as model organisms.

The fission yeast *Schizosaccharomyces pombe* is a unicellular eukaryote with a genome divided among three chromosomes [Reviewed in (Forsburg and Rhind, 2006)]. Comparison of gene sequences suggests that *S. pombe* is more similar to mammalian cells than *S. cerevisiae*. Also, its cell cycle and chromosome structure are in general more analogous to mammalian cells compared to budding yeast (Moreno et al., 1991). *S. pombe* divides by medial fission, producing two identical daughter cells. The nuclear cell cycle is divided into distinct G1 (10%), S (10%), G2 (70%) and M (10%) phases and the generation time in normal minimal or complex media is between 2 and 4 hours. *S. pombe* cells prefer a vegetative growth as haploid and may be of two mating types: *h+* and *h−* (Moreno et al., 1991). Starving induces the haploid cells of opposite mating types to mate in pairs, forming diploid zygotes, which then undergo meiosis to form four haploid spores in a linear tetrad ascus (Forsburg and Rhind, 2006). As in other organisms, meiotic recombination in fission yeast is initiated by Spo-11/Rec12 (De Veaux et al., 1992; Keeney et al., 1997).

The soil nematode *Caenorhabditis elegans* is a widely used model organism for animal development, because of its simplicity as a multicellular organism, its rapid (3-day) life cycle, and ease to knock down gene expression by RNA interference (Riddle, 1997). Furthermore, it has a well characterized germline development, where different stages of meiosis can be visualized and compared in the same organism simultaneously. The two sexes of *C. elegans*, hermaphrodites and males, are of similar length but differ in appearance as adults - males have a slim torso, move rapidly and have a characteristic hook-shaped tail (Riddle, 1997). Hermaphrodites produce both oocytes and sperm and can reproduce by self-fertilisation. Each worm lays up to 300 eggs which hatch in 12 hours, producing larvae (Brenner, 1974). The nuclear genome of *C. elegans* is organized into six chromosomes (Brenner, 1974). The karyotype of *C. elegans* consists of 12 chromosomes in hermaphrodites, five pairs of autosomes and two sex chromosomes, and 11 chromosomes in males, which have a single sex chromosome. In *C. elegans*, X0 males arise spontaneously in XX hermaphrodite populations by means of X chromosome nondisjunction at a frequency of about

0.1% (Riddle, 1997). The males mate with hermaphrodites to produce a 1:1 ratio of male and hermaphrodite cross-progeny. The chromosomes of *C. elegans* undergo all of the classically described stages of meiotic prophase, culminating in a reductional division at meiosis I and equational division at meiosis II. As in *S. pombe*, homologous chromosomes in *C. elegans* rely on the formation of crossovers to ensure their proper disjunction at meiosis I, however in the nematode, due to crossover interference, these are limited to one crossover per chromosome pair (Villeneuve, 1994).

4.3.2 Results.

4.3.2.1 Generation of expression vectors for fission yeast studies.

For analysis of base mismatches in fission yeast meiosis, we generated vectors expressing the indicated DNA deaminases and an AID catalytic mutant (Figure 46. For details see 3.2.5.3). AID catalytic mutatant possesses a mutation in the zinc co-ordinating amino acid histidine 56 (H56L) (Chaudhuri et al., 2003; Morgan et al., 2004; Papavasiliou and Schatz, 2002; Pham et al., 2003).
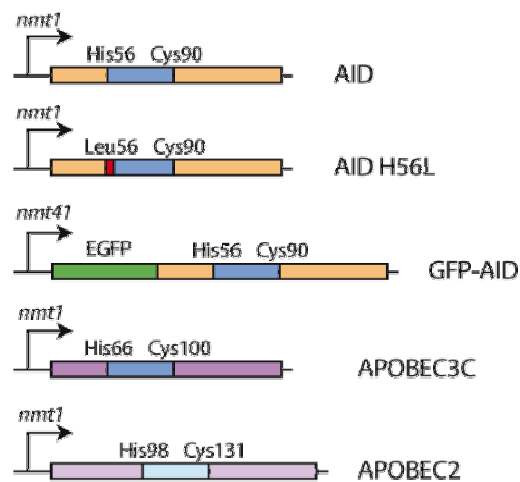
Figure 40. Schematic representation of the constructs used in fission yeast study. Proteins were expressed under the control of thiamine-repressible *nmt*1 or *nmt*41 promoter - expression is driven in the absence of thiamine. Amino acids essential for the formation of the catalytic domain (blue) are indicated as well as the substitution of His56 with Leu56 in human AID (AID).

4.3.2.2 Induction of base mismatches rescues tetrad formation in *S. pombe* *rec12*Δ.

In *S. pombe*, meiotic recombination is initiated by the formation of DSBs induced by the Spo-11 equivalent in fission yeast - Rec12 (De Veaux et al., 1992). *Rec12*Δ mutants are deficient in tetrad formation (Figure 47A) and therefore meiosis yields only approximately 50 % of tetrads (compared to 94 % in *wt*) and 30 % dyads (1-2 % dyads in *wt*), indicative of impaired meiotic recombination. The expression of AID during meiosis in the *rec12*Δ strain significantly rescued tetrad formation and decreased dyads (Figure 47B). Tetrad formation was

increased from 50 % to 73 %, while dyad formation decreased from 30 % to 12 %, indicating that AID or its dU product can compensate for some of the functions of Rec12 (Pauklin et al., 2009a). If AID expression was repressed in these strains by addition of thiamine, the tetrad rescue was abolished (Table 11). To determine, if the catalytic activity of AID is required for *rec12Δ* rescue, we expressed a catalytically inactive mutant of AID during meiosis (Figure 46). Expression of AID H56L in the *rec12Δ* background failed to rescue the deficiency in tetrad formation (Figure 47C). To follow protein expression *in vivo*, we used a GFP-AID fusion protein (Figure 46). Similarly to functional AID, this fusion protein was also able to partially rescue tetrad formation, at the same time lowering the number of dyads (Figure 47D and Table 11).
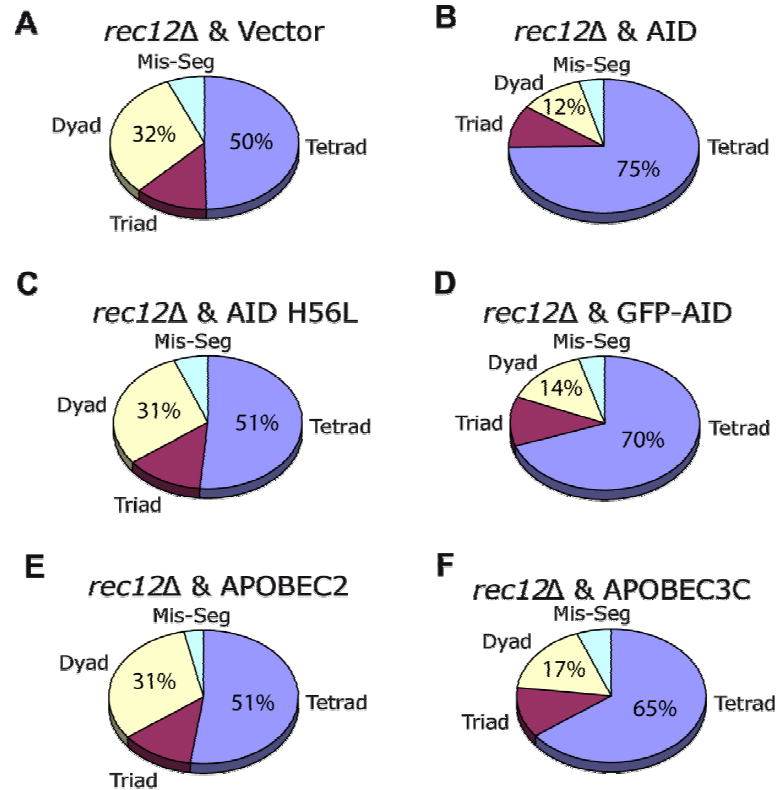
Figure 41. DNA deaminases partially restore tetrad formation in *rec12Δ S. pombe. (A - F) Relative distribution of tetrads, triads, dyads and mis-segregations, that are formed during meiosis in *rec12Δ* background transfected with an empty vector (A), with human AID (B), with a H56L human AID (C), with GFP-human AID (D), with human APOBEC2 (E), and with human APOBEC3C (F). Cells were transfected, grown in the presence of thiamine for 24 hours, washed, grown in the absence of thiamine for 6 hours, and crossed for 2 days on thiamine (-) plates. Tetrads, triads, dyads and mis-segregations were quantitated by microscopy. Modified from (Pauklin et al., 2009a).

Table 11. Asci analysis in *S. pombe*. All vectors were expressed in *rec12Δ S. pombe*. The concentration of thiamine was 15 µM. The numbers highlighted in red correspond to the tetrad and dyad percentages in AID-expressing cells. n = number of asci. Published in (Pauklin et al., 2009a).

% of analysed asci	Tetrad		Triad		Dyad		Mis-seg		n	
vector \ thiamine	-	+	-	+	-	+	-	+	-	+
pREP1	49.8	50.1	12.5	12.2	31.2	31.5	6.5	6.2	514	520
pREP1 AID (H56L)	51.4	50.9	13.2	13.6	29.7	30.2	5.7	5.3	508	512
pREP1 AID	**74.6**	52.4	10.5	11.8	**10.7**	29.3	4.2	6.5	536	502
pREP41 GFP-AID	69.5	51.2	11.8	12.2	14.3	29.9	4.4	6.7	525	542
pREP1 APOBEC2	50.6	50.2	12.3	12.6	30.8	31.2	6.3	6.0	506	558
pREP1 APOBEC3C	64.8	51.3	12.1	12.7	17.2	30.8	5.9	5.2	542	523

4.3.2.3 AID expression in *rec12Δ* increases spore viability.

DNA deaminases generate uracil in DNA, followed by the processing of these lesions by BER to abasic sites or ssDNA nicks, which in turn can be processed to DSBs. In our tetrad analysis, this could have increased DSB lesions to an extent that started decreasing spore viabilities. We addressed this by dissecting tetrads, triads, and dyads into individual spores, to analyse spore viabilities. The average spore viability from a tetrad of *rec12Δ* was 22 %, but when catalytically active AID was expressed during meiosis, viability increased by more than 2 fold (Figure 48A). An increase in viability was also observed when spores were picked from dyads or triads (Figure 48A), or when a random spore analysis from cultures was preformed (Figure 48B). As in the tetrad formation assay, the catalytically inactive AID mutant was unable to rescue the viability.
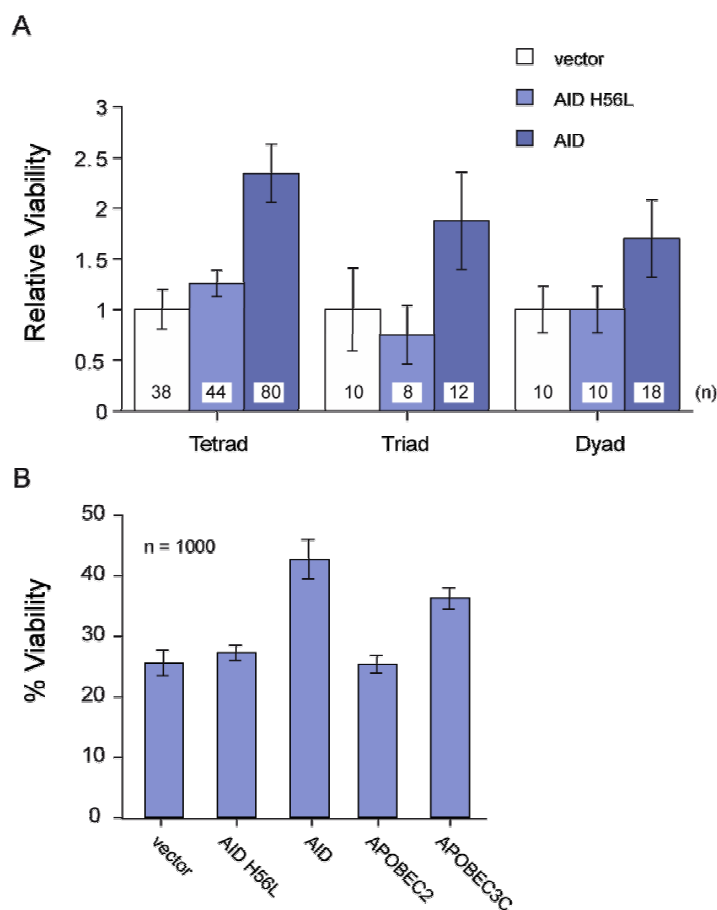
Figure 42. Deaminases increase spore viability after meiosis. (A) Relative spore viability by ascus dissection analysis. Viabilities of *rec12Δ* single spores from tetrads, triads and dyads in control-transfected cells (white), in AID catalytic mutant expressing cells (light blue) and in catalytically active AID expressing cells (dark blue). The numbers of the spores that were isolated are indicated at the bottom of each column, and the viability of vector only *rec12Δ* strain set to one. (B) Percent viability of random spore analysis, from *rec12Δ* cells expressing different deaminases during meiosis. Published in (Pauklin et al., 2009a).

4.3.2.4 AID expression induces meiotic recombination.

DNA DSBs generated by Spo-11/Rec12 induce inter-homologue recombination that promotes proper chromosome segregation by inducing chiasmata (Bahler et al., 1993; Gerton and Hawley, 2005; Whitby, 2005). To investigate whether AID expression and the formation of dU could lead to meiotic recombination, we analysed the recombination frequency between *ade6* and *ura4*, two auxotrophic markers on chromosome III. In *rec12Δ* strain, meiotic recombination frequency between these markers is approximately 0.1 - 0.2 %, a value derived from the number of viable spores growing on plates lacking uracil and adenine versus the

total spore count. Expression of AID during meiosis in *rec12Δ* strains induced a 15-fold increase in recombination (Figure 49B) (Pauklin et al., 2009a). This was dependent on the expression of AID as no recombination was detected in the presence of thiamine, as well as AID's catalytic activity, because AID H56L did not significantly increase recombination (Table 12).

Table 12. Initiation of meiotic recombination by DNA deaminases.

vector	% Recombination	Fold Change
pREP1	0.11	1.0
pREP1 AID (H56L)	0.18	1.5
pREP1 AID	**1.74**	**15.0**
pREP41 GFP-AID	1.38	12.0
pREP1 APOBEC2	0.13	1.2
pREP1 APOBEC3C	0.46	4.0

all vectors were expressed in *rec12Δ*;
% recombination = wt phenotype/ total input

Next, we analysed whether recombination derived from dU can be initiated after DNA replication. Cells were synchronized at the post-replicative stage of meiosis by using nocodazole (known to depolymerise microtubules). To monitor the efficacy of the treatment (completion of S phase) we used propidium-iodide staining and FACS analysis (Appendix Figure 6). Yeast were grown in the presence of thiamine (time-line in Figure 49C), induced to mate at time 0 h, nocodazole was added after 24 h, a further 24 h later AID was expressed by removing thiamine [expression was monitored via GFP-AID (see Appendix Figure 6)] while the cells were still in nocodazole. By 72 h post mating initiation, cells were released from the nocodazole block by extensive washing and AID expression inhibited with thiamine. After 24 h, cells were plated and incubated until colonies were visible to score for meiotic recombination. If AID was induced after the nocodazole block, a significant increase in the frequency of meiotic recombination between the two loci was still observed (Pauklin et al., 2009a). As expected, this effect required AID expression and AID catalysis (Figure 49D). These results suggest that a significant proportion of the AID-mediated meiotic recombination events occur independently from DNA replication.
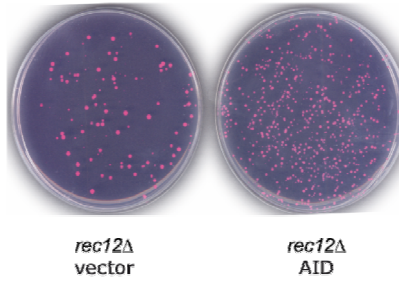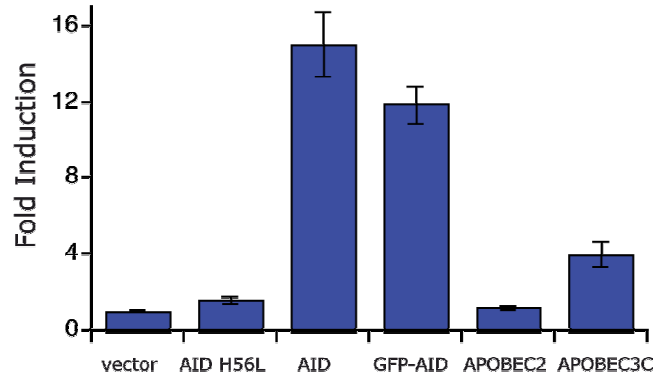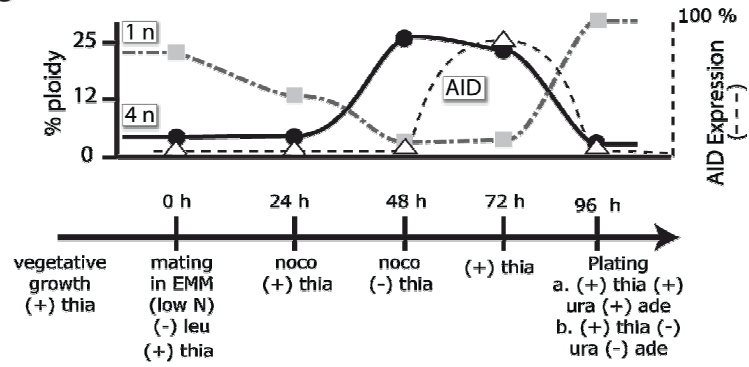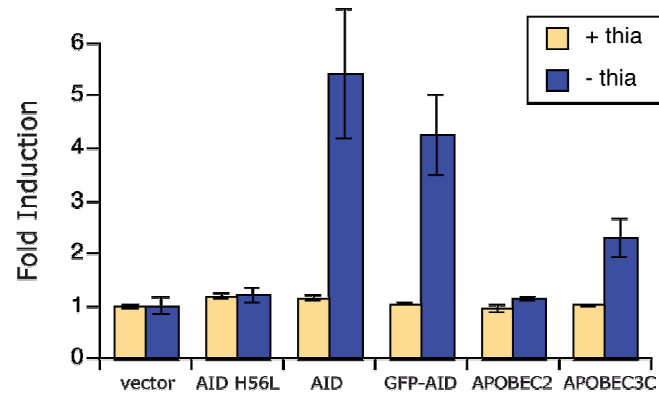
A

rec12Δ
vector

rec12Δ
AID

B

C

D

151

Figure 43. Meiotic recombination initiated by AID. (A) Representative plates showing the increase in recombinants after meiosis of *rec12Δ* cells transfected with empty vector (left plate) or with AID (right plate). Cells auxotrophic for either adenine (*ade6*) or uracil (*ura4*) were crossed on EMM low N and spores plated on EMM media supplemented with phloxin B, and in the absence or presence of both adenine and uracil, to assess the relative number of recombinants compared to the total number of viable spores. (B) Fold increase in *ade6 ura4* recombinants after the expression of either AID catalytic mutant, functional AID, GFP-AID, APOBEC2 or APOBEC3C during meiosis compared to empty vector (set to one) transfectants. (C) Timeline of post-replicative induction of AID-induced recombination. Time points on the figure represent the relative timing of changing the media or the addition of the indicated compounds nocodazole (noco), leucine (leu), and thiamine (thia); + and – indicate the presence or absence of the compounds starting from the corresponding time point. (D) DNA deaminase mediated recombination events are initiated after DNA replication. Cells were treated as indicated in the timeline (C). The graph shows the recombinants obtained when AID, mutant (H56L) AID and the APOBEC proteins were expressed after removal of thiamine (- thia) at 48 hour time point compared to the results when deaminase expression is repressed during meiosis by thiamine (+ thia). For details see 3.2.5.9. Published in (Pauklin et al., 2009a).

It is thought that the sub-cellular localization of AID and other APOBEC family members may be a means of regulating the function of DNA deaminases (Coker and Petersen-Mahrt, 2007). AID is known to enter and function in the nucleus, whereas the APOBEC3 family members (except for APOBEC3B) are usually confined to the cytoplasm (Mangeat et al., 2003; Marin et al., 2003). We therefore tested if other DNA deaminases besides AID, could be used to induce dU:dG mismatches during meiosis leading to rescue of the *rec12Δ* phenotype as well as homologous recombination. Expression vectors for APOBEC3C (a DNA deaminase known to inhibit SIV and Line1 retrotransposon (Dutko et al., 2005; Harris et al., 2002a; Muckenfuss et al., 2006; Yu et al., 2004b) and APOBEC2 (Conticello et al., 2005; Harris et al., 2002a; Jarmuz et al., 2002; Mikl et al., 2005)] were transformed into fission yeast. Expression of APOBEC2 failed to rescue tetrad formation (Figure 47E), increase viability (Figure 48B), or increase meiotic recombination (Table 12). On the other hand, expression of APOBEC3C did increase tetrad formation (Figure 47F), viability (Figure 48B), as well as meiotic recombination (Table 12), indicating that dU from other DNA deaminases can also induce meiotic recombination events.

4.3.2.5 Generating an AID expressing *spo-11* deficient nematode.

Results from the single cell fission yeast meiosis suggested, that a base mismatch

generated by AID's enzymatic activity results in the processing of the lesion leading to homologous recombination. To study the effects of AID expression in a genetically tractable multicellular organism, we generated a nematode that expresses AID in meiotic cells.

AID was fused to GFP and placed under the control of the *pie-1* promoter, limiting its expression to germline and early embryogenesis. *Pie-1* promoter was originally derived from a nematode gene known to express specifically in the germ line and during early embryogenesis (Tenenhaus et al., 1998). The *pie-1* AID vector was used for worm microparticle bombardment (described in 3.2.6), using an *unc-119* (UNCoordinated) mutant worm strain. Vector integration confers movement to the worm due to unc-119 protein expression. Upon stable integration, the progeny of this worm maintains the ability to move. Two weeks post-bombardment, plates were analysed under a light microscope, showing 10 positive plates out of 120 plates. From each of these positive plates, 6 moving worms were picked and transferred to a 24-well plate. Progeny of the worms were visually analysed under light microscopy. Animals from 8 out of 10 plates gave predominantly moving progeny, indicative of stable integration of the AID expression vector. The integration of AID sequence was confirmed on individual animals by nested PCR (Figure 50; see also 3.2.6.9), using primers for amplifying AID sequence and analysing on average 9 animals per plate.
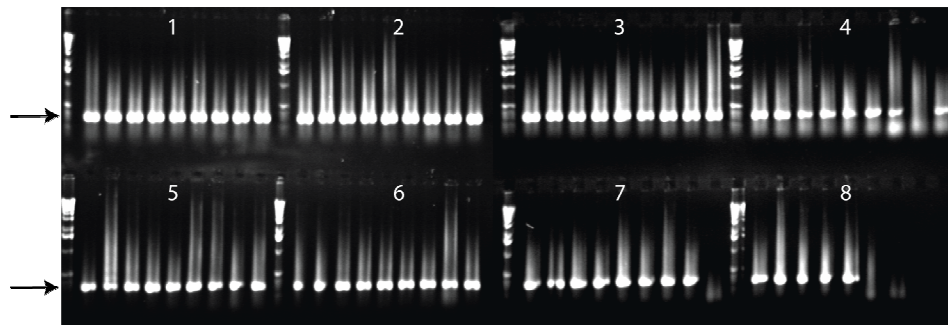
Figure 44. Confirmation of the stable integration of AID in the nematode. Moving progeny of bombarded worms were used in single worm PCR to detect the genomic integration of AID. Numbers indicate different plates from which worms were picked for analysis. Arrows mark the PCR product corresponding to AID sequence.

In addition, live worms were also analysed for GFP signal with Deltavision microscope in high magnification, indicating a weak GFP signal predominantly in the cytoplasm of germline cells. This cytoplasmic staining was confirmed with monoclonal antibodies against AID protein (Figure 52). Immunocytological staining of the animals indicated the presence of AID protein throughout the germline from meiosis I leptotene through diplotene. As predicted, AID signal was absent from *spo-11* worms.

In order to obtain a *spo-11* deficient nematode expressing AID, several consecutive crosses were carried out as schematically depicted in Figure 51. Firstly, *spo-11* males were crossed with a GFP balancing strain that harbours a
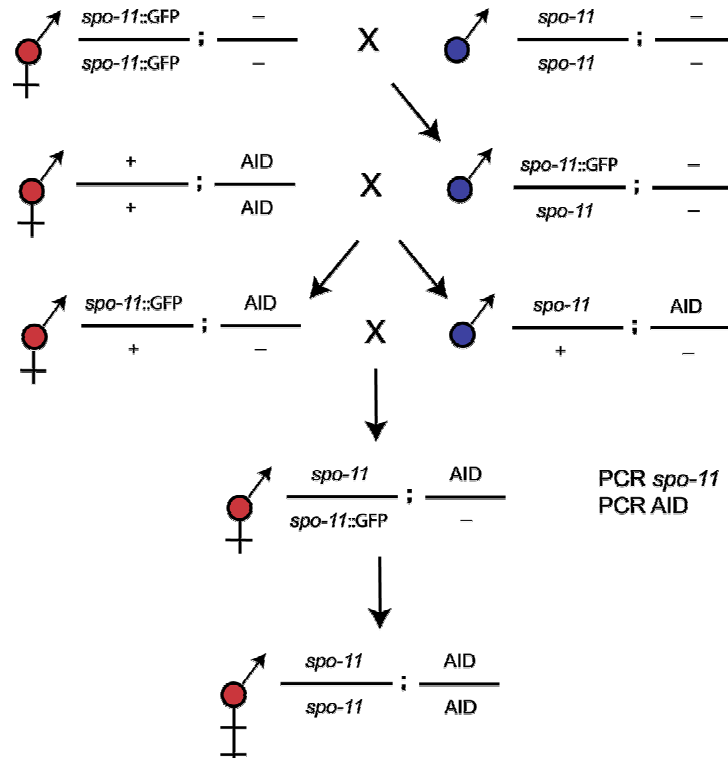
Figure 51. A schematic representation of nematode crossing. The strategy for obtaining a *spo-11* deficient nematode stably expressing AID in its germ line. *Spo-11* marks the deficient allele, GFP marks the balanced allele, "+" marks the wt *spo-11* alleele, "-" marks the absence of AID. Female and male are depicted with red and blue symbols, respectively.

GFP ORF very close to the genomic locus of *spo-11* gene. Since the *spo-11* gene

and GFP loci are in physical proximity, the GFP signal (exclusively in the pharynx of the animal) can be considered as a genetic marker for this *spo-11* allele, allowing to trace its segregation among the progeny. GFP positive male progeny (a total of 5 animals) from this cross were crossed with the above AID *unc-119* hermaphrodites (2 animals). Thereafter, GFP expressing hermaphrodites that arose from the cross, were further cross with non-GFP males from the same brood. Hermaphrodites from the resulting brood were then individually grown to give self progeny. These worms were placed in individual wells of a 24-well plate and allowed to reproduce. The animals were subjected to nested PCR to detect both AID and the deletion of *spo-11* gene. Nested PCR results allowed for the identification of PCR double-positive homozygous AID expressing *spo-11* deficient animals. These animals were propagated and used for subsequent studies.
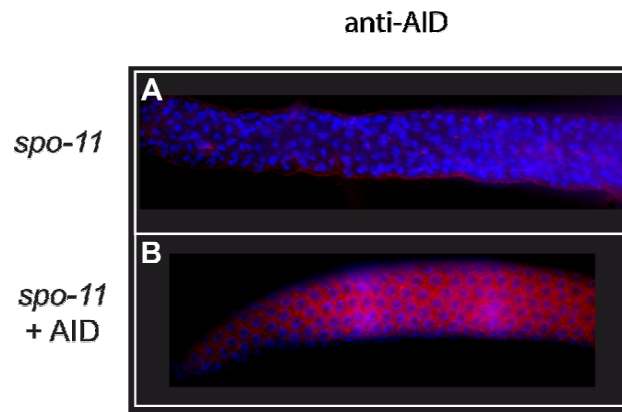
Figure 45. Expression of AID during oogenesis. AID was detected with anti-AID antibodies (red) in (A) *spo-11* worm and (B) AID expressing *spo-11* worm. Nuclei were counter stained with DAPI. Published in (Pauklin et al., 2009a).

4.3.2.6 AID expression rescues RAD-51 focus formation and bivalency in *spo-11* mutant *C. elegans*.

In *C. elegans*, each chromosome pair has a single crossover event (Villeneuve, 1994). Animals deficient in Spo-11 are devoid of meiotic RAD-51 focus formation and fail to generate the obligate crossover per homolog pair. Cytologically, this manifests as chromosome univalency at diakinesis, that can be

visualized as separate DAPI staining bodies inside the nucleus (Dernburg et al., 1998). To determine if ectopic expression of AID in the *C. elegans* germline is capable of inducing meiotic recombination, *spo-11* deficient worms expressing AID were examined for RAD-51 focus formation as a measure for meiotic recombination (Figure 53). Wild-type worms exhibit the normal distribution of Spo-11 induced RAD-51 foci, which are located predominantly in late leptotene to pachytene (Figure 53A and G). As expected, the *spo-11* mutants are deficient in the induction of meiotic DSBs and therefore lack RAD-51 foci (Figure 53B). Ectopic germline expression of AID induces meiotic RAD-51 foci (Figure 53C). The distribution of these foci is broader and less numerous than in wild-type organism, but interestingly are most abundant near the leptotene and zygotene zone (Figure 53C and G). Importantly, AID-induced RAD-51 foci were absent from the pre-meiotic zone, which contains the only actively proliferating nuclei in the adult animal, indicating that lesions induced by AID are mostly post-replicative. Our work in *S. pombe* (Figure 49D) is in agreement with this, indicating that DNA replication is dispensable for the AID induced meiotic recombination. Previous studies have used ionising radiation to artificially induce DSBs in *spo-11* deficient worm in order to rescue the *spo-11* phenotype (Dernburg et al., 1998). Gamma radiation induced RAD-51 focus formation in *spo-11* worms are distributed randomly throughout meiotic prophase (Figure 53E), which is in contrast to the AID-induced lesions, which peaked around zygotene (Figure 53C and G), suggesting differences between these sources of DNA lesions.
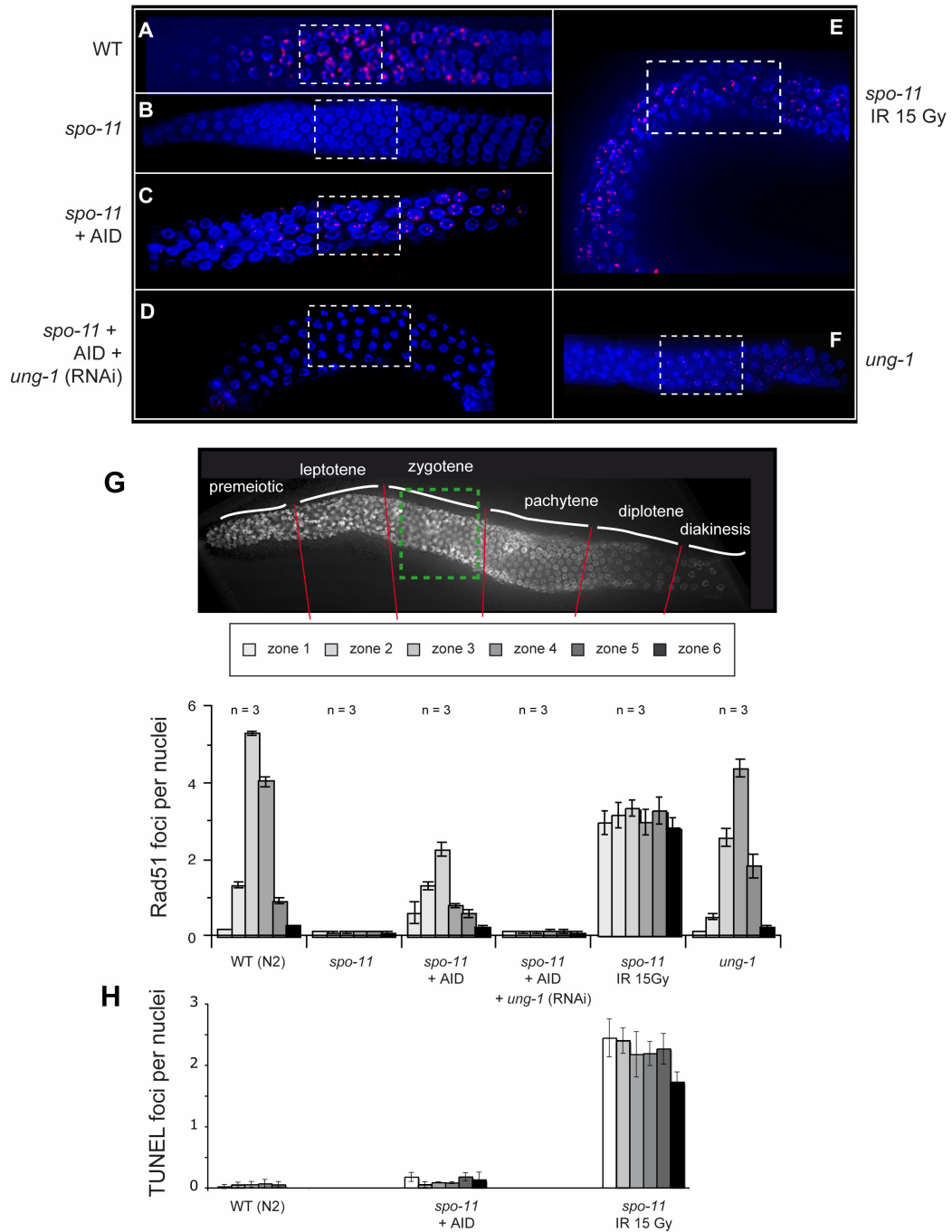
Figure 46. AID induces RAD-51 foci during meiosis in *spo-11* worm. Representative pictures of germlines visualized for RAD-51 foci, using anti-RAD-51 antibodies (red) as well as DAPI staining to visualise the nuclei: (A) wild-type animals, (B) *spo-11*, (C) *spo-11* worm expressing AID, (D) *spo-11* worm expressing AID with RNAi treatment for *ung-1*, (E) *spo-11* worm irradiated with 15 Grays (Gy) and analysed after 4 hours, and (F) *ung-1* defective nematode. Dashed squares represent the proposed location of zygotene stage in the germline. The germlines were divided into six zones according to the developmental stage of the germline. (G) The number of RAD-51 foci per nuclei was counted in worm germlines with respective genotypes (A-F) in the six developmental zones (30 nuclei per zone per worm, analyzing 3 worms per genotype). Published in (Pauklin et al., 2009a).

During meiosis, a condensed homologous chromosome pair held physically together by a single chiasma, known as a bivalent. In a *spo-11* worm, there is no manifestation of bivalency, leading to the formation of 12 univalent chromosomes (Figure 54B). In wild-type nematode, DAPI staining reveals 6 bivalents in oocyte nuclei at diakinesis (Figure 54A). Upon AID expression in *spo-11* deficient animals, some of the AID induced mismatches are processed to induce meiotic recombination leading to a substantial recovery of bivalency (Figure 54C and E), indicating that AID is able to influence chromosome segregation.
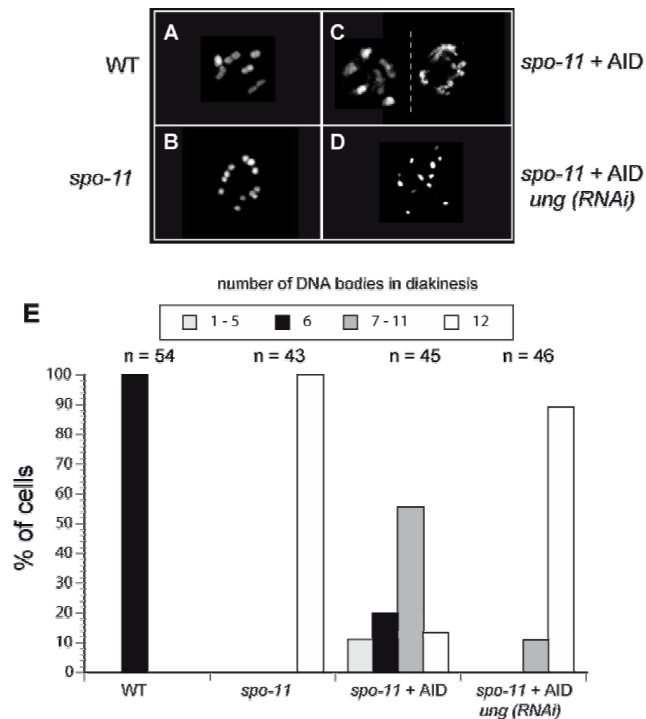
Figure 47. AID alters the chromosome morphology during diakinesis. Representative pictures of DAPI stained chromosomes during diakinesis from (A) wild-type, (B) *spo-11*, (C) two nuclei of AID expressing *spo-11* worm, and (D) AID expressing *spo-11* worm after treatment with *ung-1* (RNAi). (E) The numbers of DAPI staining bodies per nuclei were counted for each genotype from the indicated number of cells. Different shadings for the columns represent the numbers of DAPI stained bodies as indicated on the top of the graph. Published in (Pauklin et al., 2009a).

4.3.2.7 AID expression increases viability and decreases males in *spo-11* nematode.

Since the formation of RAD-51 foci and chromosomal bivalency in response to AID expression indicated changes in the molecular events taking place during meiosis compared to *spo-11* worms, we investigated if AID could also increase the viability of the *spo-11* mutant worms. Due to aberrant chromosome segregation, *spo-11* worms produce less than 1% viable offspring, compared to wild-type N2 worms (Figure 55A; Table 13). Chromosomal non-dysjunction in these animals also results in increased frequency of male offspring, which is increased from ~ 0.1% in wild-type N2 to above 40 % in *spo-11* mutants (Figure 55B and Table 12) (Dernburg et al., 1998). *C. elegans spo-11* expressing AID exhibit a more than 10 fold increase in viable brood size (Figure 55A) and a 2.5 fold decrease in males (Figure 55B and Table 13). These results indicate that the induction of meiotic recombination events by expressing AID in *spo-11* mutants is sufficient to increase viability of the progeny and proper chromosome segregation.
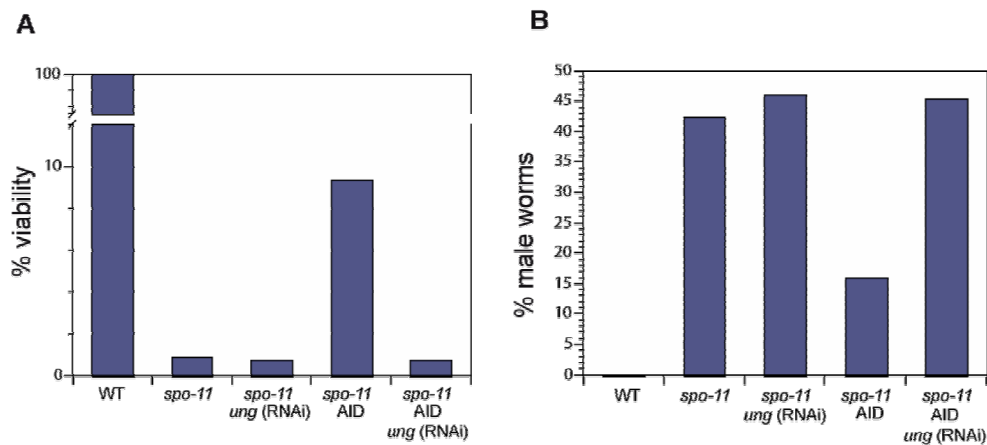
Figure 48. AID increases viability and hermaphrodite frequency in *spo-11* worms. (A) Worm viabilities for the indicated genotypes were assessed by counting viable offspring and dead eggs in a brood from young adults. (B) The frequencies of males among viable progeny from the sample in (A) were assessed by microscopic analysis of tail morphology and movement. A total of at least 2000 worms were counted per genotype (for extended data see Table 12). Published in (Pauklin et al., 2009a).

Table 13. Viability and male frequency in the nematode. Viability (% Viability) is based on the number of viable offspring/ (dead eggs + viable offspring); % Lethality = 100 - Viability; Male frequency (% Male) is based on the numbers of animals with a male phenotype (hook-sphaped tail, slim torso, more active movement) among at least 170 viable offspring; Hermaphrodite frequency (% Hermaph.) = 100 - % Male; wt - N2 or *unc-119*; si-UNG - siRNA UNG treated worms; 15 Gy - 15 Grey units of ionizing radiation; †p<0.01 (to wt);††p<0.05 (to *spo-11*); *p<0.001 (to *spo-11*); **p<0.01 (to wt si-UNG);*†p<0.01 (to *spo-11* + AID). Published in (Pauklin et al., 2009a).

Strain	Dead eggs	Live progeny	Total counts	% Lethality	% Viability	Male	Hermaph.	% Male	% Hermaph.
wt	0	4333	4333	0.0	100.0	0	4333	0.0	100.0
spo-11	4797	42	4839	99.1	0.9	151	208	42.1	57.9
wt + AID	17†	2043	2060	0.8	99.2	2	2041	0.1	99.9
spo-11 + AID	4401	440	4841	90.9	9.1*	64	376	14.5*	85.5
spo-11 + AID si-UNG	2267	13	2280	99.4	0.6*†	80	97	45.2*†	54.8
wt si-UNG	5	2027	2032	0.3	99.8	2	2025	0.1	99.9
spo-11 si-UNG	4853	25††	4878	99.5	0.5	159	184	46.4	53.6
wt + AID si-UNG	59	1962	2021	2.9**	97.1	37	1925	1.9**	98.1
ung-1	12†	2134	2146	0.6	99.4	9†	2125	0.4	99.6
wt IR 15 Gy	134	2119	2253	6.0	94.1	5	2114	0.2	99.8
spo-11 IR 15 Gy	2291	25	2316	98.9	1.1	96	163	37.1	63.9

4.3.2.8 Rescue of *spo-11* by AID is dependent on UNG.

To investigate which DNA repair pathways may be involved in processing the AID-induced dU:dG lesion, we inactivated UNG-1 [the only Uracil DNA glycosylase in nematode (Dengg et al., 2006)] by RNAi in AID expressing *spo-11* worms. UNG-1 depletion by RNAi completely inhibited AID induced RAD-51 focus formation in *spo-11* mutants (Figure 53D and G). Similarly to RAD-51 foci formation, AID induced bivalency and the decrease in male frequency was dependent on the presence of UNG, since RNAi for UNG abrogated these effects (Figure 54D and E). Furthermore, analysis of *ung-1* (RNAi) treated wt worms indicated a minor but significant increase in dead eggs and male frequencies (Table 13). These effects were even larger in the *ung-1* worm (knock-out), and also detectable in Spo-11 deficient background in which *ung-1* was depleted by RNAi (Table 13), altogether indicating that BER proteins can affect meiosis. According to these results, AID induced dU lesions seem to be predominantly processed via the BER pathway to generate abasic sites that are sufficient to generate HR substrates in meiosis.

Since AID expression resulted in RAD-51 foci that peaked around zygotene (Figure 53C), we investigated the expression of UNG-1 in the germline using

specific antibodies (Figure 56). Even though there were minor differences in the distribution of UNG-1 compared to AID expression (Figure 52 vs 56), there is no clear indication that the relative distribution of AID induced RAD-51 foci is solely caused by the expression pattern of UNG. Therefore, the distribution of these foci might be caused by other features of the meiotic recombination pathway or differences in cell stage throughout oogenesis.
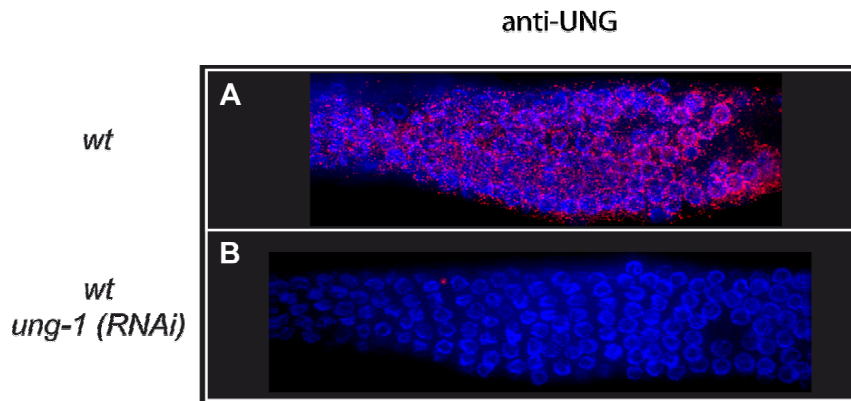
Figure 49. Expression of UNG during oogenesis. UNG was detected with anti-UNG antibody (red) in (A) wild-type worm germline and (B) wild-type worm treated with *ung-1* (RNAi). Nuclei were counterstained with DAPI. Published in (Pauklin et al., 2009a).

4.3.2.9 Differences in AID and gamma radiation induced rescue of *spo-11* phenotype.

Gamma radiation induced RAD-51 focus formation in *spo-11* worms were distributed randomly throughout meiotic prophase (Figure 53E), whereas AID-induced lesions peaked around zygotene (Figure 53C). It was possible there are differences in the processing of irradiation induced lesions and AID-induced lesions. To directly visualize the formation of DSBs and accessible DNA ends, we employed a previously published TUNEL assay (Parusel et al., 2006). Although we could observe a large number of TUNEL positive cells after gamma radiation treatment (sub-lethal dose), we did not detect foci in the AID expressing worms (Figure 57B and C). This demonstrated, that AID did not induce a large number of DSBs or accessible DNA ends unlike the ionizing radiation treatment and therefore AID induced lesions differ from lesions generated by gamma irradiation.
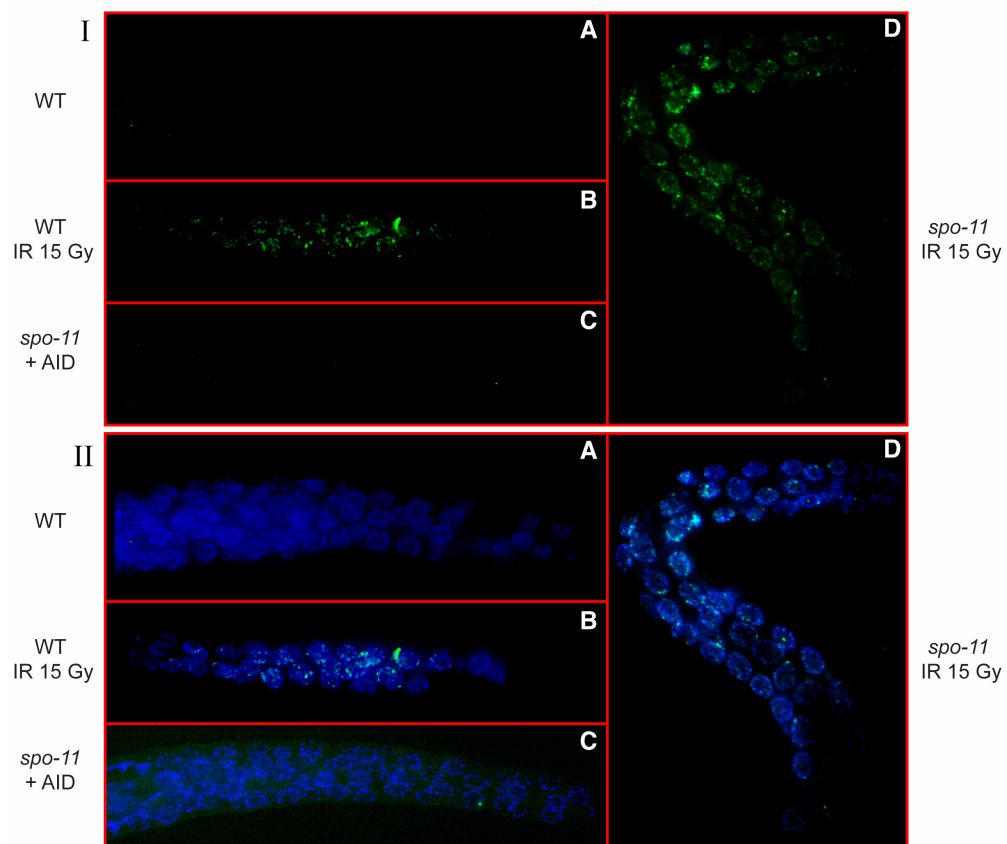
Figure 50. TUNEL staining of *C. elegans* germline. Worms were treated as described in 3.2.6.19 and 3.2.6.20. One representative germline from wt (A), wt (IR 15 Gy) (B), *spo-11* + AID (C), and *spo-11* (IR 15 Gy) (D) is shown for TUNEL staining (I - upper) and merged with DAPI (II - lower) fluorescence. Published in (Pauklin et al., 2009a).

4.3.3 Discussion.

The genome is under constant threat by chemical or physical sources of DNA damage. These lesions need to be removed by DNA repair mechanisms in order to maintain the integrity of the genome. A common single base altering lesion in non-replicative DNA is deaminated cytosine, resulting in the formation of a dU:dG mismatch. In the Ig locus, different DNA repair pathways can process the dU:dG lesion leading to various phases of SHM, iGC, and CSR (Chaudhuri and Alt, 2004; Maizels, 2005; Neuberger et al., 2005; Neuberger et al., 2003; Petersen-Mahrt et al., 2002). The repair of a dU:dG mismatch leading to class switching is mediated by BER (UNG) as well as MMR (MLH1-PMS2, Msh2/Msh6 and MSH5) (Larson et al., 2005b; Rada et al., 2004; Sekine et al., 2007), followed by NHEJ (Ku70/Ku80 DNA PKs) that completes the process (Manis et al., 1998; Rooney et al., 2004). In my study, I was interested in the fate of dU:dG lesions in meiotic cells, using AID as a source for these lesions. To date, it is not known if AID can function in a meiotic nucleus, if dU:dG mismatches are repaired in meiotic cells and if so, which DNA repair pathways are responsible for it. Since it has been hypothesized that DNA lesions other than Spo-11 induced DSBs might be involved in meiotic recombination (Cromie et al., 2006; Farah et al., 2005), we speculated that AID-induced deamination could affect these recombination events.

Expression of DNA deaminases in the *rec12Δ* strains of *S. pombe* lead to a rescue of tetrad formation (Figure 47), increased spore viability (Figure 48), and increased meiotic recombination (Figure 49). The partial rescue of *C. elegans spo-11* by AID demonstrates that in multi-cellular organisms the meiotic nucleus is also capable of processing a dU lesion. The formation of RAD-51 foci indicates that AID induced dU lesions can be processed by the post-replicative meiotic nucleus to HR substrates (Figure 53C). Furthermore, the increased chromosomal bivalency (Figure 54C), viable offspring (Figure 55A) and decrease in males (Figure 55B), suggests that AID induced base mismatches can give rise to crossovers that are capable of guiding chromosomal segregation at the first meiotic division. Our RNAi experiments in the nematode indicated that the processing of dU lesions into recombination events in meiotic cells was dependent on UNG-1, suggesting that BER influences AID-induced DNA repair

processes during meiosis.

The rescue of Spo-11 deficiency by AID in both fission yeast and nematode is important, considering the difference of the lesions that are generated by AID (dU) and Spo-11 (DSB). Since Spo-11 may have auxiliary functions during meiosis, it might explain the partial rescue of *spo-11* phenotype by AID. AID expression in the nematode germline led to RAD-51 foci, which were in modestly lower numbers than in wt, however the viability and male rescue was not complete nor to the extent to which the RAD-51 foci would have indicated (Figure 53C and Figure 55). As each chromosome pair has to acquire one recombination event (visualized by a RAD-51 focus), AID induced lesions may not have allowed for sufficient HR substrate formation. This might have only partially restored proper chromosomal segregation or even cause some chromosomal instability, as indicated by irregular bivalency numbers (Figure 54C). Future work could address whether AID lacks some auxiliary function as Spo-11 or whether the lesions generated by AID and Spo-11 undergo different processing in these cells.

Importantly, AID expression in germline did not induce RAD-51 foci in the pre-meiotic zone of the germline (Figure 53C and G), which contains the only cells in the germline that replicate their DNA. This suggests, that the majority of RAD-51 foci induced by AID in *spo-11* mutants occur in post-replicative cells. This is supported by the results in fission yeast, which also indicated that replication is not essential for AID-induced recombination (Figure 49D). It remains to be determined, if this is due to differences in the accessibility of AID to ssDNA or a difference in DNA repair between the pre-meiotic and meiotic phases. It can not be attributed to UNG-1 expression, which is expressed in both stages (Figure 56A).

In B-cells, the processing of AID induced mismatches and the formation of DSBs seem to be dependent on the genetic loci, the cell state, the chromatin status and AID co-factors, as DSBs are limited to the highly repetitive switch regions during class switch recombination, but are less likely to occur during somatic hypermutation of the variable regions (Petersen et al., 2001; Stavnezer et al., 2008; Wuerffel et al., 1997). By using TUNEL assay we analysed free DNA ends indicative of DSBs in gamma irradiated (sub-lethal dose), wt and AID expressing worms. In gamma irradiated worms, we observed extensive TUNEL

staining throughout the germline but no TUNEL staining in wt animals. Though the TUNEL assay has a detection limit, it is not clear if this lack of signal is due to the threshold of detection, or if the Spo-11 induced DSBs are inaccessible. Since Spo-11 is temporarily covalently-bound to the DSB during meiosis (Keeney et al., 1997), and can co-localise with the lesion to recruit downstream processing pathways, it might result in "hiding" the available DSB until other repair factors are present. From the TUNEL staining, there did not appear to be a large number of AID induced DSBs or free ends during meiosis. This implies, that although base mismatches are processed into recombination events as indicated by RAD-51 foci and chromosome bivalency, the lesions seem to be different in their accessibility to TUNEL staining. Thus, the lack of TUNEL staining and the enhanced efficiency of the rescued viability of AID-induced base mismatches over IR (Table 13), indicate that random DNA DSB generation is not as efficient in rescuing the *spo-11* phenotype, compared to AID induced base lesions.

My results show for the first time that DNA lesions other than DSB such as dU, abasic site, or ssDNA nicks, can initiate meiotic recombination in yeast and the multi-cellular organism *C. elegans* independently from DNA replication. This suggests an evolutionary conserved mechanism of DNA repair for dU lesions in the meiotic nucleus of organisms as diverse as fission yeast and the nematode. Since AID and some other DNA deaminases are expressed in oocytes (Morgan et al., 2004) and during spermatogenesis (Schreck et al., 2006), it is possible that protein induced DNA deamination can affect meiotic recombination also in vertebrates.

Chapter 5

CONCLUSIONS AND PROSPECTS

5. Conclusions and prospects.

5.1 Hormonal regulation of AID.

AID is essential for antibody diversification by deaminating cytosines to uracils. Due to hormonal effects on immune responses and the expression of AID in some hormone-responsive tissues, I investigated whether AID could be regulated by estrogen and progesterone. Estrogen treatment resulted in an increase in AID transcription via the binding of ER to AID promoter. In addition, estrogen could cooperate with the NFκB pathway in the regulation of AID expression. Importantly, up-regulation of AID mRNA resulted in an increase in SHM and CSR. This not only has implications for the efficiency of humoral immune responses, but also autoimmune pathologies and malignancies caused by estrogen induced AID-mediated genomic instability. In addition to the Ig loci, we found increases in mutation frequencies in the tumor suppressor CD95/Fas loci and translocations between Ig and c-myc loci.

Since estrogen resulted in an increase in AID mRNA also in ovarian tissue, future work should address the role AID plays in this tissue. One possible function could be inhibiting the movement of retro-transposable elements. This possibility in supported by our data on estrogen induced increase in Apobec3, its primary function being to inhibit retro-transposable elements and retroviral infections. Since there are considerable fluctuations in the levels of estrogen in females during the reproductive cycle (Smith et al., 2000), peaking close to the timing of fertilisation, one can speculate that estrogen mediated up-regulation of some APOBEC3 family members could provide an additional mechanism for fighting off retroviral infections.

In contrast to estrogen, progesterone resulted in a decrease in AID expression in stimulated B-cells. This was mediated by the binding of PR to AID promoter, resulting in the attenuation of SHM and CSR. These results place AID in the regulatory pathways of estrogen and progesterone that modulate the down-stream processes of AID by regulating its transcription.

5.2 Identification of AID-interacting proteins.

By generating four DT40 cell lines with endogenously tagged AID, I established a system that allowed me to investigate AID localisation and to identify novel proteins that form a complex with AID in the cytoplasm, nucleoplasm and on chromatin. My investigation of the sub-cellular distribution of AID is in agreement with earlier studies that most of the total AID protein pool is localised in the cytoplasm, and a small fraction resides in the nucleoplasm and on the chromatin. This distribution could be explained by the need of keeping AID out of the nucleus to avoid potentially harmful mutations in the genome. Indeed, using an AID-eGFP DT40 cell line, I found that the endogenous AID protein is actively exported from the nucleus via a CRM1-dependent export mechanism, correlating with other studies. However, it is likely that this is not the only mechanism regulating the intracellular distribution of AID.

By using large scale IP and subsequent mass spectrometry, I was able to identify a number of proteins interacting with AID. Since this approach indicated several proteins in the cytoplasmic fraction and nucleoplasm that could potentially be involved in the nucleo-cytoplasmic transport of AID, it would be important to study, whether these proteins bind AID directly. AID has been suggested to be regulated by post-translational modifications. Since AID has been shown to be phosphorylated by PKC, future studies should address the question, whether PKC adaptor protein RACK1 or other signaling proteins identified in my study, regulate AID phosphorylation and its function.

My mass spectrometry results in the chromatin fraction suggest that AID resides in a complex with elongating RNA polymerase. This could modulate AID-mediated cytosine deamination and help to explain the distribution of SHM in Ig loci. Since AID has been shown to bind directly to the splicing factor CTNNBL1, it would be useful to identify the transcripts that are bound to AID complex in the chromatin fraction, to find out whether these RNA species are spliced in a specific way or contain specific sequence motifs. This might explain, why only a fraction of transcribed loci are targeted by AID.

The DT40 cell lines generated in this study will be useful for analysing the regulation of AID and post-translational modifications in the future. My investigation provides an essential framework for understanding the process of

DNA deamination *in vivo* and for investigating these novel AID-interacting proteins in more detail.

5.3 Dissecting DNA repair pathways downstream of AID-induced deamination.

AID expression during meiosis in *spo-11/rec12* deficient fission yeast and nematode was able to partially rescue a meiotic defect phenotype. The appearance of RAD-51 foci in AID expressing *spo-11* nematode suggested that lesions generated by AID were processed into meiotic recombination events. This was confirmed by increased meiotic crossovers and the visualisation of bivalent homologous chromosomes. Interestingly, AID expression also resulted in abnormal chromosomal structures in some nuclei, suggesting that AID can cause chromosomal re-arrangements. Importantly, in both yeast and worm, AID was able to initiate meiotic recombination independently from DNA replication. Therefore, uracils in DNA can be processed by the meiotic nucleus to give rise to meiotic recombination. In worm, this was mediated by BER since knockdown of UNG by RNAi abolished AID-induced *spo-11* rescue. These results raise the possibility, that recombination triggered by U lesions might also occur in mammalian germ cells.

Since errors in chromosomal segregation are correlated with unusual crossover positions in many cases of maternally derived trisomy 21 (Lamb et al., 1997; Lamb et al., 1996), it is possible that spontaneous DNA deamination or AID-induced uracils can interfere with Spo-11 induced crossovers, or give rise to crossovers outside of normal recombination "hot-spots". Analogously, increased aneuploidy has been noted for irradiation (Neel, 1998). In the light of my finding that AID mRNA is up-regulated in the ovarian tissue in response to estrogen and present in oocytes, future studies could address the question, whether AID expression in germ cells can affect the frequency of chromosomal abnormalities and aneuploidy. This could be addressed in AID knockout and wt mice treated with estrogen, and analysed for the frequencies of aneuploidy.

The deamination of cytosine to uracil is central for the biological function of AID. This deamination event is not only essential for humoral immunity, but it is involved in pathologies – autoimmunity and oncogenesis. Therefore it is important to understand, how AID is regulated, both on the level of gene expression, and its binding partners in the cell. Once AID has generated the dU lesion, it is essential to delineate, what are the mechanisms that determine, whether this lesion undergoes error free repair or is utilized for the distinct DNA repair pathways involved in SHM, iGC or CSR. The results from the three diverse approaches described in my thesis, have all addressed questions that are central for understanding AID-mediated DNA deamination: 1) how is AID expression regulated in the cell? 2) which factors interact with AID in different subcellular compartments, and modulate AID function? 3) what determines the fate of the dU lesion generated by AID? The results of my studies have significantly furthered the general understanding of the process of AID-mediated DNA deamination, and opened up new ways for addressing these important and exciting questions in the future.

References

Abbas, A.K., Lichtmann, A.H., Pillai, S. (2007). Cellular and Molecular Immunology.

Allen, C.D., Okada, T., and Cyster, J.G. (2007). Germinal-center organization and cellular dynamics. Immunity *27*, 190-202.

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

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

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

Arakawa, H., Moldovan, G.L., Saribasak, H., Saribasak, N.N., Jentsch, S., and Buerstedde, J.M. (2006). A role for PCNA ubiquitination in immunoglobulin hypermutation. PLoS Biol *4*, e366.

Arakawa, H., Saribasak, H., and Buerstedde, J.M. (2004). Activation-induced cytidine deaminase initiates immunoglobulin gene conversion and hypermutation by a common intermediate. PLoS Biol *2*, E179.

Arnalich, F., Benito-Urbina, S., Gonzalez-Gancedo, P., Iglesias, E., de Miguel, E., and Gijon-Banos, J. (1992). Inadequate production of progesterone in women with systemic lupus erythematosus. Br J Rheumatol *31*, 247-251.

Auboeuf, D., Honig, A., Berget, S.M., and O'Malley, B.W. (2002). Coordinate regulation of transcription and splicing by steroid receptor coregulators. Science *298*, 416-419.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (1987). Current Protocols in Molecular Biology (Boston, JohnWiley & Sons, Inc.).

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

Bahler, J., Wyler, T., Loidl, J., and Kohli, J. (1993). Unusual nuclear structures in meiotic prophase of fission yeast: a cytological analysis. J Cell Biol *121*, 241-256.

Barnes, D.E., and Lindahl, T. (2004). Repair and genetic consequences of endogenous DNA base damage in mammalian cells. Annu Rev Genet *38*, 445-476.

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

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

Baudat, F., Manova, K., Yuen, J.P., Jasin, M., and Keeney, S. (2000). Chromosome synapsis defects and sexually dimorphic meiotic progression in mice lacking Spo11. Mol Cell *6*, 989-998.

Beato, M., and Klug, J. (2000). Steroid hormone receptors: an update. Hum Reprod Update *6*, 225-236.

Bebo, B.F., Jr., Fyfe-Johnson, A., Adlard, K., Beam, A.G., Vandenbark, A.A., and Offner, H. (2001). Low-dose estrogen therapy ameliorates experimental autoimmune encephalomyelitis in two different inbred mouse strains. J Immunol *166*, 2080-2089.

Bebo, B.F., Jr., Schuster, J.C., Vandenbark, A.A., and Offner, H. (1999). Androgens alter the cytokine profile and reduce encephalitogenicity of myelin-reactive T cells. J Immunol *162*, 35-40.

Bennett, R.A., Wilson, D.M., 3rd, Wong, D., and Demple, B. (1997). Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway. Proc Natl Acad Sci U S A *94*, 7166-7169.

Benten, W.P., Stephan, C., and Wunderlich, F. (2002). B cells express intracellular but not surface receptors for testosterone and estradiol. Steroids *67*, 647-654.

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

Bishop, K.N., Holmes, R.K., and Malim, M.H. (2006). Antiviral potency of APOBEC proteins does not correlate with cytidine deamination. J Virol *80*, 8450-8458.

Bjornstrom, L., and Sjoberg, M. (2005). Mechanisms of estrogen receptor signaling: convergence of genomic and nongenomic actions on target genes. Mol Endocrinol *19*, 833-842.

Both, G.W., Taylor, L., Pollard, J.W., and Steele, E.J. (1990). Distribution of mutations around rearranged heavy-chain antibody variable-region genes. Mol Cell Biol *10*, 5187-5196.

Boulton, S.J., Gartner, A., Reboul, J., Vaglio, P., Dyson, N., Hill, D.E., and Vidal, M. (2002). Combined functional genomic maps of the C. elegans DNA damage response. Science *295*, 127-131.

Bransteitter, R., Pham, P., Scharff, M.D., and Goodman, M.F. (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.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71-94.

Buerstedde, J.-M., and Takeda, S. (2006). Reviews and Protocols in DT40 Research (New York, Springer).

Butterworth, M., McClellan, B., and Allansmith, M. (1967). Influence of sex in immunoglobulin levels. Nature *214*, 1224-1225.

Casellas, R., Nussenzweig, A., Wuerffel, R., Pelanda, R., Reichlin, A., Suh, H., Qin, X.F., Besmer, E., Kenter, A., Rajewsky, K.*, et al.* (1998). Ku80 is required for immunoglobulin isotype switching. EMBO J *17*, 2404-2411.

Cerillo, G., Rees, A., Manchanda, N., Reilly, C., Brogan, I., White, A., and Needham, M. (1998). The oestrogen receptor regulates NFkappaB and AP-1 activity in a cell-specific manner. J Steroid Biochem Mol Biol *67*, 79-88.

Cervera, R., Khamashta, M.A., Font, J., Sebastiani, G.D., Gil, A., Lavilla, P., Domenech, I., Aydintug, A.O., Jedryka-Goral, A., de Ramon, E.*, et al.* (1993). Systemic lupus erythematosus: clinical and immunologic patterns of disease expression in a cohort of 1,000 patients. The European Working

Party on Systemic Lupus Erythematosus. Medicine (Baltimore) *72*, 113-124.

Chaudhuri, J., and Alt, F.W. (2004). Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. Nat Rev Immunol *4*, 541-552.

Chaudhuri, J., Khuong, C., and Alt, F.W. (2004). Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. Nature *430*, 992-998.

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

Chen, J.M., Cooper, D.N., Chuzhanova, N., Ferec, C., and Patrinos, G.P. (2007). Gene conversion: mechanisms, evolution and human disease. Nat Rev Genet *8*, 762-775.

Chesi, M., Robbiani, D.F., Sebag, M., Chng, W.J., Affer, M., Tiedemann, R., Valdez, R., Palmer, S.E., Haas, S.S., Stewart, A.K.*, et al.* (2008). AID-dependent activation of a MYC transgene induces multiple myeloma in a conditional mouse model of post-germinal center malignancies. Cancer Cell *13*, 167-180.

Chiu, Y.L., and Greene, W.C. (2008). The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. Annu Rev Immunol *26*, 317-353.

Chiu, Y.L., Witkowska, H.E., Hall, S.C., Santiago, M., Soros, V.B., Esnault, C., Heidmann, T., and Greene, W.C. (2006). High-molecular-mass APOBEC3G complexes restrict Alu retrotransposition. Proc Natl Acad Sci U S A *103*, 15588-15593.

Christen, U., and von Herrath, M.G. (2004). Initiation of autoimmunity. Curr Opin Immunol *16*, 759-767.

Cohen-Solal, J.F., Jeganathan, V., Grimaldi, C.M., Peeva, E., and Diamond, B. (2006). Sex hormones and SLE: influencing the fate of autoreactive B cells. Curr Top Microbiol Immunol *305*, 67-88.

Coker, H.A., and Petersen-Mahrt, S.K. (2007). The nuclear DNA deaminase AID functions distributively whereas cytoplasmic APOBEC3G has a processive mode of action. DNA Repair (Amst) *6*, 235-243.

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

Conticello, S.G., Langlois, M.A., Yang, Z., and Neuberger, M.S. (2007). DNA deamination in immunity: AID in the context of its APOBEC relatives. Adv Immunol *94*, 37-73.

Conticello, S.G., Thomas, C.J., Petersen-Mahrt, S.K., and Neuberger, M.S. (2005). Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. Molecular biology and evolution *22*, 367-377.

Cromie, G.A., Hyppa, R.W., Taylor, A.F., Zakharyevich, K., Hunter, N., and Smith, G.R. (2006). Single Holliday Junctions Are Intermediates of Meiotic Recombination. Cell *127*, 1167-1178.

Davidson, A., and Diamond, B. (2001). Autoimmune diseases. N Engl J Med *345*, 340-350.

De Veaux, L.C., Hoagland, N.A., and Smith, G.R. (1992). Seventeen complementation groups of mutations decreasing meiotic recombination in Schizosaccharomyces pombe. Genetics *130*, 251-262.

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

Dedeoglu, F., Horwitz, B., Chaudhuri, J., Alt, F.W., and Geha, R.S. (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.

Delbos, F., Aoufouchi, S., Faili, A., Weill, J.C., and Reynaud, C.A. (2007). DNA polymerase eta is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. J Exp Med *204*, 17-23.

Delbos, F., De Smet, A., Faili, A., Aoufouchi, S., Weill, J.C., and Reynaud, C.A. (2005). Contribution of DNA polymerase eta to immunoglobulin gene hypermutation in the mouse. J Exp Med *201*, 1191-1196.

Demple, B., and Sung, J.S. (2005). Molecular and biological roles of Ape1 protein in mammalian base excision repair. DNA Repair (Amst) *4*, 1442-1449.

Dengg, M., Garcia-Muse, T., Gill, S.G., Ashcroft, N., Boulton, S.J., and Nilsen, H. (2006). Abrogation of the CLK-2 checkpoint leads to tolerance to base-excision repair intermediates. EMBO Rep *7*, 1046-1051.

Dernburg, A.F., McDonald, K., Moulder, G., Barstead, R., Dresser, M., and Villeneuve, A.M. (1998). Meiotic recombination in C. elegans initiates by a conserved mechanism and is dispensable for homologous chromosome synapsis. Cell *94*, 387-398.

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

Dickerson, S.K., Market, E., Besmer, E., and Papavasiliou, F.N. (2003). AID mediates hypermutation by deaminating single stranded DNA. J Exp Med *197*, 1291-1296.

Dorner, T., and Radbruch, A. (2007). Antibodies and B cell memory in viral immunity. Immunity *27*, 384-392.

Dorsett, Y., McBride, K.M., Jankovic, M., Gazumyan, A., Thai, T.H., Robbiani, D.F., Di Virgilio, M., San-Martin, B.R., Heidkamp, G., Schwickert, T.A.*, et al.* (2008). MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. Immunity *28*, 630-638.

Druckmann, R., and Druckmann, M.A. (2005). Progesterone and the immunology of pregnancy. J Steroid Biochem Mol Biol *97*, 389-396.

Dunnick, W., Hertz, G.Z., Scappino, L., and Gritzmacher, C. (1993). DNA sequences at immunoglobulin switch region recombination sites. Nucleic Acids Res *21*, 365-372.

Duquette, M.L., Pham, P., Goodman, M.F., and Maizels, N. (2005). AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation. Oncogene *24*, 5791-5798.

Durandy, A., Peron, S., Taubenheim, N., and Fischer, A. (2006). Activation-induced cytidine deaminase: structure-function relationship as based on the study of mutants. Hum Mutat *27*, 1185-1191.

Dutko, J.A., Schafer, A., Kenny, A.E., Cullen, B.R., and Curcio, M.J. (2005). Inhibition of a yeast LTR retrotransposon by human APOBEC3 cytidine deaminases. Curr Biol *15*, 661-666.

Eidinger, D., and Garrett, T.J. (1972). Studies of the regulatory effects of the sex hormones on antibody formation and stem cell differentiation. J Exp Med *136*, 1098-1116.

Endo, Y., Marusawa, H., Kinoshita, K., Morisawa, T., Sakurai, T., Okazaki, I.M., Watashi, K., Shimotohno, K., Honjo, T., and Chiba, T. (2007). Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. Oncogene *26*, 5587-5595.

Ermann, J., and Fathman, C.G. (2001). Autoimmune diseases: genes, bugs and failed regulation. Nat Immunol *2*, 759-761.

Esnault, C., Heidmann, O., Delebecque, F., Dewannieux, M., Ribet, D., Hance, A.J., Heidmann, T., and Schwartz, O. (2005). APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. Nature *433*, 430-433.

Farah, J.A., Cromie, G., Davis, L., Steiner, W.W., and Smith, G.R. (2005). Activation of an alternative, rec12 (spo11)-independent pathway of fission yeast meiotic recombination in the absence of a DNA flap endonuclease. Genetics *171*, 1499-1511.

Felix, N.J., and Allen, P.M. (2007). Specificity of T-cell alloreactivity. Nat Rev Immunol *7*, 942-953.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Font, J., Pallares, L., Cervera, R., Lopez-Soto, A., Navarro, M., Bosch, X., and Ingelmo, M. (1991). Systemic lupus erythematosus in the elderly: clinical and immunological characteristics. Ann Rheum Dis *50*, 702-705.

Forsburg, S.L., and Rhind, N. (2006). Basic methods for fission yeast. Yeast *23*, 173-183.

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

Gallois-Montbrun, S., Kramer, B., Swanson, C.M., Byers, H., Lynham, S., Ward, M., and Malim, M.H. (2007). Antiviral protein APOBEC3G localizes to ribonucleoprotein complexes found in P bodies and stress granules. J Virol *81*, 2165-2178.

Gerton, J.L., and Hawley, R.S. (2005). Homologous chromosome interactions in meiosis: diversity amidst conservation. Nat Rev Genet *6*, 477-487.

Ghosh, M.G., Thompson, D.A., and Weigel, R.J. (2000). PDZK1 and GREB1 are estrogen-regulated genes expressed in hormone-responsive breast cancer. Cancer Res *60*, 6367-6375.

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

Goodnow, C.C. (2007). Multistep pathogenesis of autoimmune disease. Cell *130*, 25-35.

Goossens, T., Klein, U., and Kuppers, R. (1998). Frequent occurrence of deletions and duplications during somatic hypermutation: implications for oncogene translocations and heavy chain disease. Proc Natl Acad Sci U S A *95*, 2463-2468.

Gordon, M.S., Kanegai, C.M., Doerr, J.R., and Wall, R. (2003). Somatic hypermutation of the B cell receptor genes B29 (Igbeta, CD79b) and mb1 (Igalpha, CD79a). Proc Natl Acad Sci U S A *100*, 4126-4131.

Grimaldi, C.M., Hicks, R., and Diamond, B. (2005). B cell selection and susceptibility to autoimmunity. J Immunol *174*, 1775-1781.

Gruber, C.J., Tschugguel, W., Schneeberger, C., and Huber, J.C. (2002). Production and actions of estrogens. N Engl J Med *346*, 340-352.

Harris, R.S., Bishop, K.N., Sheehy, A.M., Craig, H.M., Petersen-Mahrt, S.K., Watt, I.N., Neuberger, M.S., and Malim, M.H. (2003). DNA deamination mediates innate immunity to retroviral infection. Cell *113*, 803-809.

Harris, R.S., and Liddament, M.T. (2004). Retroviral restriction by APOBEC proteins. Nat Rev Immunol *4*, 868-877.

Harris, R.S., Petersen-Mahrt, S.K., and Neuberger, M.S. (2002a). RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. Mol Cell *10*, 1247-1253.

Harris, R.S., Sale, J.E., Petersen-Mahrt, S.K., and Neuberger, M.S. (2002b). AID is essential for immunoglobulin V gene conversion in a cultured B cell line. Curr Biol *12*, 435-438.

Holmes, R.K., Koning, F.A., Bishop, K.N., and Malim, M.H. (2007a). APOBEC3F can inhibit the accumulation of HIV-1 reverse transcription products in the absence of hypermutation. Comparisons with APOBEC3G. J Biol Chem *282*, 2587-2595.

Holmes, R.K., Malim, M.H., and Bishop, K.N. (2007b). APOBEC-mediated viral restriction: not simply editing? Trends Biochem Sci *32*, 118-128.

Hsu, H.C., Wu, Y., Yang, P., Wu, Q., Job, G., Chen, J., Wang, J., Accavitti-Loper, M.A., Grizzle, W.E., Carter, R.H.*, et al.* (2007). Overexpression of activation-induced cytidine deaminase in B cells is associated with production of highly pathogenic autoantibodies. J Immunol *178*, 5357-5365.

Ing, N.H., Beekman, J.M., Tsai, S.Y., Tsai, M.J., and O'Malley, B.W. (1992). Members of the steroid hormone receptor superfamily interact with TFIIB (S300-II). J Biol Chem *267*, 17617-17623.

Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M., and Honjo, T. (2004). Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. Proc Natl Acad Sci U S A *101*, 1975-1980.

Janeway, C. (2001). Immunobiology.

Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J., and Navaratnam, N. (2002). An anthropoid-specific locus of orphan C to U RNA-editing enzymes on chromosome 22. Genomics *79*, 285-296.

Jiang, C., Foley, J., Clayton, N., Kissling, G., Jokinen, M., Herbert, R., and Diaz, M. (2007). Abrogation of lupus nephritis in activation-induced deaminase-deficient MRL/lpr mice. J Immunol *178*, 7422-7431.

Kalkhoven, E., Wissink, S., van der Saag, P.T., and van der Burg, B. (1996). Negative interaction between the RelA(p65) subunit of NF-kappaB and the progesterone receptor. J Biol Chem *271*, 6217-6224.

Kamradt, T., and Mitchison, N.A. (2001). Tolerance and autoimmunity. N Engl J Med *344*, 655-664.

Kauppi, L., Jeffreys, A.J., and Keeney, S. (2004). Where the crossovers are: recombination distributions in mammals. Nat Rev Genet *5*, 413-424.

Keeney, S., Giroux, C.N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell *88*, 375-384.

Kinoshita, K., Harigai, M., Fagarasan, S., Muramatsu, M., and Honjo, T. (2001). A hallmark of active class switch recombination: transcripts directed by I promoters on looped-out circular DNAs. Proc Natl Acad Sci U S A *98*, 12620-12623.

Klapholz, S., Waddell, C.S., and Esposito, R.E. (1985). The role of the SPO11 gene in meiotic recombination in yeast. Genetics *110*, 187-216.

Klein, U., Tu, Y., Stolovitzky, G.A., Keller, J.L., Haddad, J., Jr., Miljkovic, V., Cattoretti, G., Califano, A., and Dalla-Favera, R. (2003). Transcriptional analysis of the B cell germinal center reaction. Proc Natl Acad Sci U S A *100*, 2639-2644.

Kovalchuk, A.L., Muller, J.R., and Janz, S. (1997). Deletional remodeling of c-myc-deregulating chromosomal translocations. Oncogene *15*, 2369-2377.

Kozak, S.L., Marin, M., Rose, K.M., Bystrom, C., and Kabat, D. (2006). The anti-HIV-1 editing enzyme APOBEC3G binds HIV-1 RNA and messenger RNAs that shuttle between polysomes and stress granules. J Biol Chem *281*, 29105-29119.

Kristjuhan, A., and Svejstrup, J.Q. (2004). Evidence for distinct mechanisms facilitating transcript elongation through chromatin in vivo. EMBO J *23*, 4243-4252.

Krogh, B.O., and Symington, L.S. (2004). Recombination proteins in yeast. Annu Rev Genet *38*, 233-271.

Krokan, H.E., Drablos, F., and Slupphaug, G. (2002). Uracil in DNA--occurrence, consequences and repair. Oncogene *21*, 8935-8948.

Kuppers, R., and Dalla-Favera, R. (2001). Mechanisms of chromosomal translocations in B cell lymphomas. Oncogene *20*, 5580-5594.

Kushner, P.J., Agard, D.A., Greene, G.L., Scanlan, T.S., Shiau, A.K., Uht, R.M., and Webb, P. (2000). Estrogen receptor pathways to AP-1. J Steroid Biochem Mol Biol *74*, 311-317.

Lahita, R.G., Bradlow, H.L., Kunkel, H.G., and Fishman, J. (1979). Alterations of estrogen metabolism in systemic lupus erythematosus. Arthritis Rheum *22*, 1195-1198.

Lamb, N.E., Feingold, E., Savage, A., Avramopoulos, D., Freeman, S., Gu, Y., Hallberg, A., Hersey, J., Karadima, G., Pettay, D.*, et al.* (1997). Characterization of susceptible chiasma configurations that increase the risk for maternal nondisjunction of chromosome 21. Hum Mol Genet *6*, 1391-1399.

Lamb, N.E., Freeman, S.B., Savage-Austin, A., Pettay, D., Taft, L., Hersey, J., Gu, Y., Shen, J., Saker, D., May, K.M.*, et al.* (1996). Susceptible chiasmate configurations of chromosome 21 predispose to non-disjunction in both maternal meiosis I and meiosis II. Nat Genet *14*, 400-405.

Larijani, M., Frieder, D., Basit, W., and Martin, A. (2005). The mutation spectrum of purified AID is similar to the mutability index in Ramos cells and in ung(-/-)msh2(-/-) mice. Immunogenetics *56*, 840-845.

Larson, E.D., Cummings, W.J., Bednarski, D.W., and Maizels, N. (2005a). MRE11/RAD50 cleaves DNA in the AID/UNG-dependent pathway of immunoglobulin gene diversification. Mol Cell *20*, 367-375.

Larson, E.D., Duquette, M.L., Cummings, W.J., Streiff, R.J., and Maizels, N. (2005b). MutSalpha binds to and promotes synapsis of transcriptionally activated immunoglobulin switch regions. Curr Biol *15*, 470-474.

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

Lee, C.H., Melchers, M., Wang, H., Torrey, T.A., Slota, R., Qi, C.F., Kim, J.Y., Lugar, P., Kong, H.J., Farrington, L.*, et al.* (2006). Regulation of the germinal center gene program by interferon (IFN) regulatory factor 8/IFN consensus sequence-binding protein. J Exp Med *203*, 63-72.

Leonhardt, S.A., Boonyaratanakornkit, V., and Edwards, D.P. (2003). Progesterone receptor transcription and non-transcription signaling mechanisms. Steroids *68*, 761-770.

Li, X., Lonard, D.M., and O'Malley, B.W. (2004). A contemporary understanding of progesterone receptor function. Mech Ageing Dev *125*, 669-678.

Liao, W., Hong, S.H., Chan, B.H., Rudolph, F.B., Clark, S.C., and Chan, L. (1999). APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. Biochem Biophys Res Commun *260*, 398-404.

Lin, Y., and Smith, G.R. (1994). Transient, meiosis-induced expression of the rec6 and rec12 genes of Schizosaccharomyces pombe. Genetics *136*, 769-779.

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. Nature *362*, 709-715.

Lindahl, T., and Wood, R.D. (1999). Quality control by DNA repair. Science *286*, 1897-1905.

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

Longerich, S., Basu, U., Alt, F., and Storb, U. (2006). AID in somatic hypermutation and class switch recombination. Curr Opin Immunol *18*, 164-174.

Lumsden, J.M., McCarty, T., Petiniot, L.K., Shen, R., Barlow, C., Wynn, T.A., Morse, H.C., 3rd, Gearhart, P.J., Wynshaw-Boris, A., Max, E.E.*, et al.* (2004). Immunoglobulin class switch recombination is impaired in Atm-deficient mice. J Exp Med *200*, 1111-1121.

MacDuff, D.A., Neuberger, M.S., and Harris, R.S. (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.

Maizels, N. (2005). Immunoglobulin gene diversification. Annu Rev Genet *39*, 23-46.

Mangeat, B., Turelli, P., Caron, G., Friedli, M., Perrin, L., and Trono, D. (2003). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. Nature *424*, 99-103.

Manis, J.P., Gu, Y., Lansford, R., Sonoda, E., Ferrini, R., Davidson, L., Rajewsky, K., and Alt, F.W. (1998). Ku70 is required for late B cell development and immunoglobulin heavy chain class switching. J Exp Med *187*, 2081-2089.

Manis, J.P., Morales, J.C., Xia, Z., Kutok, J.L., Alt, F.W., and Carpenter, P.B. (2004). 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. Nat Immunol *5*, 481-487.

Marin, M., Rose, K.M., Kozak, S.L., and Kabat, D. (2003). HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. Nat Med *9*, 1398-1403.

Martin, A., Bardwell, P.D., Woo, C.J., Fan, M., Shulman, M.J., and Scharff, M.D. (2002). Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. Nature *415*, 802-806.

Martin, J.S., Winkelmann, N., Petalcorin, M.I., McIlwraith, M.J., and Boulton, S.J. (2005). RAD-51-dependent and -independent roles of a Caenorhabditis elegans BRCA2-related protein during DNA double-strand break repair. Mol Cell Biol *25*, 3127-3139.

Matsumoto, Y., Marusawa, H., Kinoshita, K., Endo, Y., Kou, T., Morisawa, T., Azuma, T., Okazaki, I.M., Honjo, T., and Chiba, T. (2007). Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. Nat Med *13*, 470-476.

Maundrell, K. (1990). nmt1 of fission yeast. A highly transcribed gene completely repressed by thiamine. J Biol Chem *265*, 10857-10864.

Maundrell, K. (1993). Thiamine-repressible expression vectors pREP and pRIP for fission yeast. Gene *123*, 127-130.

Mayorov, V.I., Rogozin, I.B., Adkison, L.R., Frahm, C., Kunkel, T.A., and Pavlov, Y.I. (2005). Expression of human AID in yeast induces mutations in context similar to the context of somatic hypermutation at G-C pairs in immunoglobulin genes. BMC Immunol *6*, 10.

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

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

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

McHeyzer-Williams, L.J., Malherbe, L.P., and McHeyzer-Williams, M.G. (2006). Checkpoints in memory B-cell evolution. Immunol Rev *211*, 255-268.

McKinnon, P.J., and Caldecott, K.W. (2007). DNA strand break repair and human genetic disease. Annu Rev Genomics Hum Genet *8*, 37-55.

McMurray, R., Keisler, D., Kanuckel, K., Izui, S., and Walker, S.E. (1991). Prolactin influences autoimmune disease activity in the female B/W mouse. J Immunol *147*, 3780-3787.

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

Mikl, M.C., Watt, I.N., Lu, M., Reik, W., Davies, S.L., Neuberger, M.S., and Rada, C. (2005). Mice deficient in APOBEC2 and APOBEC3. Mol Cell Biol *25*, 7270-7277.

Mol, C.D., Izumi, T., Mitra, S., and Tainer, J.A. (2000). DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. Nature *403*, 451-456.

Moreno, S., Klar, A., and Nurse, P. (1991). Molecular genetic analysis of fission yeast Schizosaccharomyces pombe. Methods Enzymol *194*, 795-823.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. J Biol Chem *279*, 52353-52360.

Muckenfuss, H., Hamdorf, M., Held, U., Perkovic, M., Lower, J., Cichutek, K., Flory, E., Schumann, G.G., and Munk, C. (2006). APOBEC3 proteins inhibit human LINE-1 retrotransposition. J Biol Chem *281*, 22161-22172.

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

Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O., and Honjo, T. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J Biol Chem *274*, 18470-18476.

Muschen, M., Re, D., Jungnickel, B., Diehl, V., Rajewsky, K., and Kuppers, R. (2000). Somatic mutation of the CD95 gene in human B cells as a side-effect of the germinal center reaction. J Exp Med *192*, 1833-1840.

Nakahara, M., Sonoda, E., Nojima, K., Sale, J.E., Takenaka, K., Kikuchi, K., Taniguchi, Y., Nakamura, K., Sumitomo, Y., Bree, R.T.*, et al.* (2009). Genetic evidence for single-strand lesions initiating Nbs1-dependent homologous recombination in diversification of Ig v in chicken B lymphocytes. PLoS Genet *5*, e1000356.

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

Navaratnam, N., Morrison, J.R., Bhattacharya, S., Patel, D., Funahashi, T., Giannoni, F., Teng, B.B., Davidson, N.O., and Scott, J. (1993). The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. J Biol Chem *268*, 20709-20712.

Neel, J.V. (1998). Genetic studies at the Atomic Bomb Casualty Commission-Radiation Effects Research Foundation: 1946-1997. Proc Natl Acad Sci U S A *95*, 5432-5436.

Neuberger, M.S., Di Noia, J.M., Beale, R.C., Williams, G.T., Yang, Z., and Rada, C. (2005). Somatic hypermutation at A.T pairs: polymerase error versus dUTP incorporation. Nat Rev Immunol *5*, 171-178.

Neuberger, M.S., Harris, R.S., Di Noia, J., and Petersen-Mahrt, S.K. (2003). Immunity through DNA deamination. Trends Biochem Sci *28*, 305-312.

Newman, E.N., Holmes, R.K., Craig, H.M., Klein, K.C., Lingappa, J.R., Malim, M.H., and Sheehy, A.M. (2005). Antiviral function of APOBEC3G can be dissociated from cytidine deaminase activity. Curr Biol *15*, 166-170.

Nonaka, T., Doi, T., Toyoshima, T., Muramatsu, M., Honjo, T., and Kinoshita, K. (2009). Carboxy-terminal domain of AID required for its mRNA complex formation in vivo. Proc Natl Acad Sci U S A *106*, 2747-2751.

Odegard, V.H., and Schatz, D.G. (2006). Targeting of somatic hypermutation. Nat Rev Immunol *6*, 573-583.

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

Okazaki, I.M., Kinoshita, K., Muramatsu, M., Yoshikawa, K., and Honjo, T. (2002). The AID enzyme induces class switch recombination in fibroblasts. Nature *416*, 340-345.

Onate, S.A., Tsai, S.Y., Tsai, M.J., and O'Malley, B.W. (1995). Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. Science *270*, 1354-1357.

Page, S.L., and Hawley, R.S. (2003). Chromosome choreography: the meiotic ballet. Science *301*, 785-789.

Papavasiliou, F.N., and Schatz, D.G. (2002). The activation-induced deaminase functions in a postcleavage step of the somatic hypermutation process. J Exp Med *195*, 1193-1198.

Parsa, J.Y., Basit, W., Wang, C.L., Gommerman, J.L., Carlyle, J.R., and Martin, A. (2007). AID mutates a non-immunoglobulin transgene independent of chromosomal position. Mol Immunol *44*, 567-575.

Parusel, C.T., Kritikou, E.A., Hengartner, M.O., Krek, W., and Gotta, M. (2006). URI-1 is required for DNA stability in C. elegans. Development *133*, 621-629.

Pasqualucci, L., Bhagat, G., Jankovic, M., Compagno, M., Smith, P., Muramatsu, M., Honjo, T., Morse, H.C., 3rd, Nussenzweig, M.C., and Dalla-Favera, R. (2008). AID is required for germinal center-derived lymphomagenesis. Nat Genet *40*, 108-112.

Pasqualucci, L., Kitaura, Y., Gu, H., and Dalla-Favera, R. (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.

Pasqualucci, L., Migliazza, A., Fracchiolla, N., William, C., Neri, A., Baldini, L., Chaganti, R.S., Klein, U., Kuppers, R., Rajewsky, K.*, et al.* (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.

Pasqualucci, L., Neumeister, P., Goossens, T., Nanjangud, G., Chaganti, R.S., Kuppers, R., and Dalla-Favera, R. (2001). Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. Nature *412*, 341-346.

Patenaude, A.M., Orthwein, A., Hu, Y., Campo, V.A., Kavli, B., Buschiazzo, A., and Di Noia, J.M. (2009). Active nuclear import and cytoplasmic retention of activation-induced deaminase. Nat Struct Mol Biol *16*, 517-527.

Pauklin, S., Burkert, J.S., Martin, J., Osman, F., Weller, S., Boulton, S.J., Whitby, M.C., and Petersen-Mahrt, S.K. (2009a). Alternative induction of meiotic recombination from single-base lesions of DNA deaminases. Genetics *182*, 41-54.

Pauklin, S., and Petersen-Mahrt, S.K. (2009). Progesterone inhibits AID by binding to the promoter. Journal of Immunology *183(2)*, 1238-44.

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

Peeva, E., Venkatesh, J., and Diamond, B. (2005). Tamoxifen blocks estrogen-induced B cell maturation but not survival. J Immunol *175*, 1415-1423.

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

Perlot, T., Li, G., and Alt, F.W. (2008). Antisense transcripts from immunoglobulin heavy-chain locus V(D)J and switch regions. Proc Natl Acad Sci U S A *105*, 3843-3848.

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

Petersen-Mahrt, S. (2005). DNA deamination in immunity. Immunol Rev *203*, 80-97.

Petersen-Mahrt, S.K., Harris, R.S., and Neuberger, M.S. (2002). AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. Nature *418*, 99-103.

Pham, P., Bransteitter, R., Petruska, J., and Goodman, M.F. (2003). Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. Nature *424*, 103-107.

Praitis, V., Casey, E., Collar, D., and Austin, J. (2001). Creation of low-copy integrated transgenic lines in Caenorhabditis elegans. Genetics *157*, 1217-1226.

Rabbitts, T.H., Forster, A., Hamlyn, P., and Baer, R. (1984). Effect of somatic mutation within translocated c-myc genes in Burkitt's lymphoma. Nature *309*, 592-597.

Rada, C., Di Noia, J.M., and Neuberger, M.S. (2004). Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. Mol Cell *16*, 163-171.

Rada, C., and Milstein, C. (2001). The intrinsic hypermutability of antibody heavy and light chain genes decays exponentially. EMBO J *20*, 4570-4576.

Rada, C., Williams, G.T., Nilsen, H., Barnes, D.E., Lindahl, T., and Neuberger, M.S. (2002). Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. Curr Biol *12*, 1748-1755.

Ramiro, A.R., Jankovic, M., Callen, E., Difilippantonio, S., Chen, H.T., McBride, K.M., Eisenreich, T.R., Chen, J., Dickins, R.A., Lowe, S.W.*, et al.* (2006). Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. Nature *440*, 105-109.

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

Ray, A., Prefontaine, K.E., and Ray, P. (1994). Down-modulation of interleukin-6 gene expression by 17 beta-estradiol in the absence of high affinity DNA binding by the estrogen receptor. J Biol Chem *269*, 12940-12946.

Reina-San-Martin, B., Chen, H.T., Nussenzweig, A., and Nussenzweig, M.C. (2004). ATM is required for efficient recombination between immunoglobulin switch regions. J Exp Med *200*, 1103-1110.

Reina-San-Martin, B., Difilippantonio, S., Hanitsch, L., Masilamani, R.F., Nussenzweig, A., and Nussenzweig, M.C. (2003). H2AX is required for

recombination between immunoglobulin switch regions but not for intra-switch region recombination or somatic hypermutation. J Exp Med *197*, 1767-1778.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A*., et al.* (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell *102*, 565-575.

Reynaud, C.A., Anquez, V., Dahan, A., and Weill, J.C. (1985). A single rearrangement event generates most of the chicken immunoglobulin light chain diversity. Cell *40*, 283-291.

Reynaud, C.A., Anquez, V., Grimal, H., and Weill, J.C. (1987). A hyperconversion mechanism generates the chicken light chain preimmune repertoire. Cell *48*, 379-388.

Riddle, B., Meyer, Priess (1997). C. elegans II.

Rieux-Laucat, F., Le Deist, F., and Fischer, A. (2003). Autoimmune lymphoproliferative syndromes: genetic defects of apoptosis pathways. Cell Death Differ *10*, 124-133.

Rieux-Laucat, F., Le Deist, F., Hivroz, C., Roberts, I.A., Debatin, K.M., Fischer, A., and de Villartay, J.P. (1995). Mutations in Fas associated with human lymphoproliferative syndrome and autoimmunity. Science *268*, 1347-1349.

Robbiani, D.F., Bothmer, A., Callen, E., Reina-San-Martin, B., Dorsett, Y., Difilippantonio, S., Bolland, D.J., Chen, H.T., Corcoran, A.E., Nussenzweig, A*., et al.* (2008). AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. Cell *135*, 1028-1038.

Rogozin, I.B., Basu, M.K., Jordan, I.K., Pavlov, Y.I., and Koonin, E.V. (2005). APOBEC4, a new member of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases predicted by computational analysis. Cell Cycle *4*, 1281-1285.

Romanienko, P.J., and Camerini-Otero, R.D. (2000). The mouse Spo11 gene is required for meiotic chromosome synapsis. Mol Cell *6*, 975-987.

Ronai, D., Iglesias-Ussel, M.D., Fan, M., Li, Z., Martin, A., and Scharff, M.D. (2007). Detection of chromatin-associated single-stranded DNA in regions targeted for somatic hypermutation. J Exp Med *204*, 181-190.

Rooney, S., Chaudhuri, J., and Alt, F.W. (2004). The role of the non-homologous end-joining pathway in lymphocyte development. Immunol Rev *200*, 115-131.

Roubinian, J., Talal, N., Siiteri, P.K., and Sadakian, J.A. (1979). Sex hormone modulation of autoimmunity in NZB/NZW mice. Arthritis Rheum *22*, 1162-1169.

Roy, D., and Liehr, J.G. (1999). Estrogen, DNA damage and mutations. Mutat Res *424*, 107-115.

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

Sale, J.E., and Neuberger, M.S. (1998). TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B cell line. Immunity *9*, 859-869.

Sanchez-Guerrero, J., Liang, M.H., Karlson, E.W., Hunter, D.J., and Colditz, G.A. (1995). Postmenopausal estrogen therapy and the risk for developing systemic lupus erythematosus. Ann Intern Med *122*, 430-433.

Saville, B., Wormke, M., Wang, F., Nguyen, T., Enmark, E., Kuiper, G., Gustafsson, J.A., and Safe, S. (2000). Ligand-, cell-, and estrogen receptor subtype (alpha/beta)-dependent activation at GC-rich (Sp1) promoter elements. J Biol Chem *275*, 5379-5387.

Sayegh, C.E., Quong, M.W., Agata, Y., and Murre, C. (2003). E-proteins directly regulate expression of activation-induced deaminase in mature B cells. Nat Immunol *4*, 586-593.

Schreck, S., Buettner, M., Kremmer, E., Bogdan, M., Herbst, H., and Niedobitek, G. (2006). Activation-induced cytidine deaminase (AID) is expressed in normal spermatogenesis but only infrequently in testicular germ cell tumours. J Pathol *210*, 26-31.

Sekine, H., Ferreira, R.C., Pan-Hammarstrom, Q., Graham, R.R., Ziemba, B., de Vries, S.S., Liu, J., Hippen, K., Koeuth, T., Ortmann, W.*, et al.* (2007). Role for Msh5 in the regulation of Ig class switch recombination. Proc Natl Acad Sci U S A.

Shapiro, G.S., and Wysocki, L.J. (2002). DNA target motifs of somatic mutagenesis in antibody genes. Crit Rev Immunol *22*, 183-200.

Shapiro-Shelef, M., Lin, K.I., Savitsky, D., Liao, J., and Calame, K. (2005). Blimp-1 is required for maintenance of long-lived plasma cells in the bone marrow. J Exp Med *202*, 1471-1476.

Sheehy, A.M., Gaddis, N.C., Choi, J.D., and Malim, M.H. (2002). Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature *418*, 646-650.

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

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

Shinkura, R., Tian, M., Smith, M., Chua, K., Fujiwara, Y., and Alt, F.W. (2003). The influence of transcriptional orientation on endogenous switch region function. Nat Immunol *4*, 435-441.

Shinohara, A., Gasior, S., Ogawa, T., Kleckner, N., and Bishop, D.K. (1997). Saccharomyces cerevisiae recA homologues RAD51 and DMC1 have both distinct and overlapping roles in meiotic recombination. Genes Cells *2*, 615-629.

Shinohara, A., Ogawa, H., and Ogawa, T. (1992). Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein. Cell *69*, 457-470.

Smith, S.M., Baskin, G.B., and Marx, P.A. (2000). Estrogen protects against vaginal transmission of simian immunodeficiency virus. J Infect Dis *182*, 708-715.

Sonoda, J., Pei, L., and Evans, R.M. (2008). Nuclear receptors: decoding metabolic disease. FEBS Lett *582*, 2-9.

Stavnezer, J., Guikema, J.E., and Schrader, C.E. (2008). Mechanism and regulation of class switch recombination. Annual review of immunology *26*, 261-292.

Stenglein, M.D., and Harris, R.S. (2006). APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. J Biol Chem *281*, 16837-16841.

Sthoeger, Z.M., Dayan, M., Tcherniack, A., Green, L., Toledo, S., Segal, R., Elkayam, O., and Mozes, E. (2003). Modulation of autoreactive responses of peripheral blood lymphocytes of patients with systemic lupus erythematosus by peptides based on human and murine anti-DNA autoantibodies. Clin Exp Immunol *131*, 385-392.

Ta, V.T., Nagaoka, H., Catalan, N., Durandy, A., Fischer, A., Imai, K., Nonoyama, S., Tashiro, J., Ikegawa, M., Ito, S.*, et al.* (2003). AID mutant analyses indicate requirement for class-switch-specific cofactors. Nat Immunol *4*, 843-848.

Taneja, V., and David, C.S. (2001). Lessons from animal models for human autoimmune diseases. Nat Immunol *2*, 781-784.

Tenenhaus, C., Schubert, C., and Seydoux, G. (1998). Genetic requirements for PIE-1 localization and inhibition of gene expression in the embryonic germ lineage of Caenorhabditis elegans. Dev Biol *200*, 212-224.

Teng, B., Burant, C.F., and Davidson, N.O. (1993). Molecular cloning of an apolipoprotein B messenger RNA editing protein. Science *260*, 1816-1819.

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

Tornwall, J., Carey, A.B., Fox, R.I., and Fox, H.S. (1999). Estrogen in autoimmunity: expression of estrogen receptors in thymic and autoimmune T cells. J Gend Specif Med *2*, 33-40.

Tsai, M.J., and O'Malley, B.W. (1994). Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Annu Rev Biochem *63*, 451-486.

Tucker, L.B., Menon, S., Schaller, J.G., and Isenberg, D.A. (1995). Adult- and childhood-onset systemic lupus erythematosus: a comparison of onset, clinical features, serology, and outcome. Br J Rheumatol *34*, 866-872.

van der Stoep, N., Gorman, J.R., and Alt, F.W. (1998). Reevaluation of 3'Ekappa function in stage- and lineage-specific rearrangement and somatic hypermutation. Immunity *8*, 743-750.

Villeneuve, A.M. (1994). A cis-acting locus that promotes crossing over between X chromosomes in Caenorhabditis elegans. Genetics *136*, 887-902.

Voegel, J.J., Heine, M.J., Zechel, C., Chambon, P., and Gronemeyer, H. (1996). TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors. EMBO J *15*, 3667-3675.

Wang, C.L., Harper, R.A., and Wabl, M. (2004). Genome-wide somatic hypermutation. Proc Natl Acad Sci U S A *101*, 7352-7356.

Ward, I.M., Reina-San-Martin, B., Olaru, A., Minn, K., Tamada, K., Lau, J.S., Cascalho, M., Chen, L., Nussenzweig, A., Livak, F.*, et al.* (2004). 53BP1 is required for class switch recombination. J Cell Biol *165*, 459-464.

Weigert, M.G., Cesari, I.M., Yonkovich, S.J., Cohn, M. (1970) Variability in the lambda light chain sequences of mouse antibody. Nature *228*,1045-7.

Weinstein, Y., Ran, S., and Segal, S. (1984). Sex-associated differences in the regulation of immune responses controlled by the MHC of the mouse. J Immunol *132*, 656-661.

Whitacre, C.C. (2001). Sex differences in autoimmune disease. Nat Immunol *2*, 777-780.

Whitby, M.C. (2005). Making crossovers during meiosis. Biochem Soc Trans *33*, 1451-1455.

Wilson, D.M., 3rd, and Barsky, D. (2001). The major human abasic endonuclease: formation, consequences and repair of abasic lesions in DNA. Mutat Res *485*, 283-307.

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

Wu, X., Darce, J.R., Chang, S.K., Nowakowski, G.S., and Jelinek, D.F. (2008). Alternative splicing regulates activation-induced cytidine deaminase (AID): implications for suppression of AID mutagenic activity in normal and malignant B cells. Blood *112*, 4675-4682.

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

Wuerffel, R.A., Du, J., Thompson, R.J., and Kenter, A.L. (1997). Ig Sgamma3 DNA-specifc double strand breaks are induced in mitogen-activated B cells and are implicated in switch recombination. J Immunol *159*, 4139-4144.

Xu, Z., Pone, E.J., Al-Qahtani, A., Park, S.R., Zan, H., and Casali, P. (2007). Regulation of aicda expression and AID activity: relevance to somatic hypermutation and class switch DNA recombination. Crit Rev Immunol *27*, 367-397.

Xue, K., Rada, C., and Neuberger, M.S. (2006). The in vivo pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in msh2-/- ung-/- mice. J Exp Med *203*, 2085-2094.

Yager, J.D., and Davidson, N.E. (2006). Estrogen carcinogenesis in breast cancer. N Engl J Med *354*, 270-282.

Yoshikawa, K., Okazaki, I.M., Eto, T., Kinoshita, K., Muramatsu, M., Nagaoka, H., and Honjo, T. (2002). AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. Science *296*, 2033-2036.

Yu, K., Huang, F.T., and Lieber, M.R. (2004a). DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine. J Biol Chem *279*, 6496-6500.

Yu, Q., Chen, D., Konig, R., Mariani, R., Unutmaz, D., and Landau, N.R. (2004b). APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. J Biol Chem *279*, 53379-53386.

Yu, X., Yu, Y., Liu, B., Luo, K., Kong, W., Mao, P., and Yu, X.F. (2003). Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. Science *302*, 1056-1060.
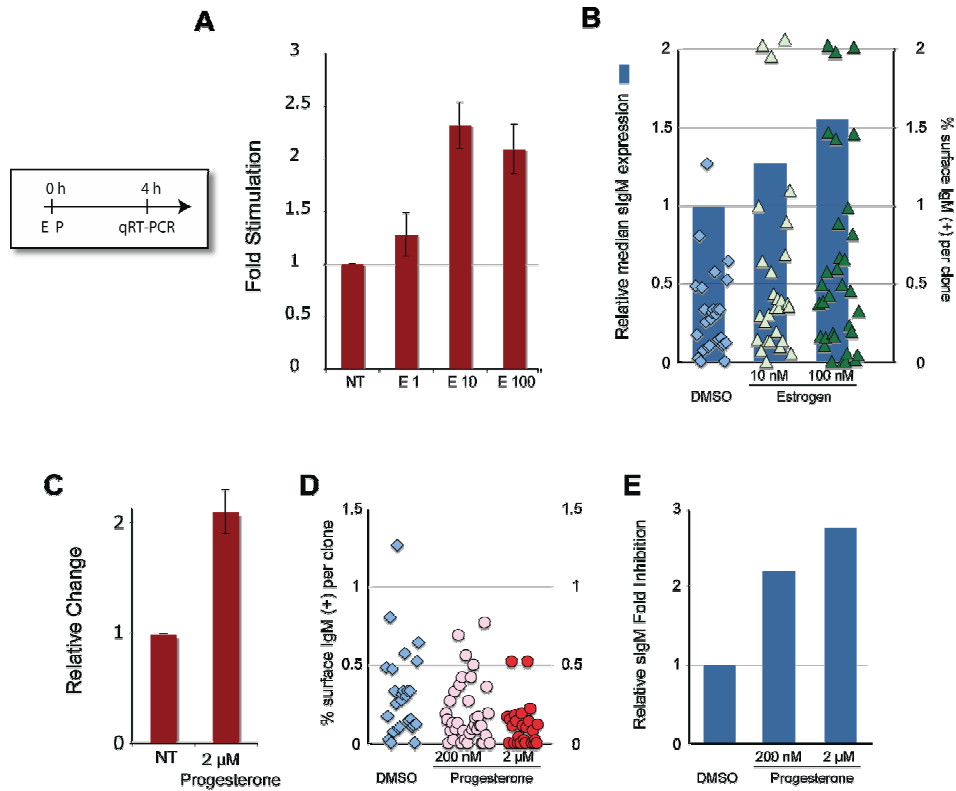
Appendix

Appendix Figure 1.
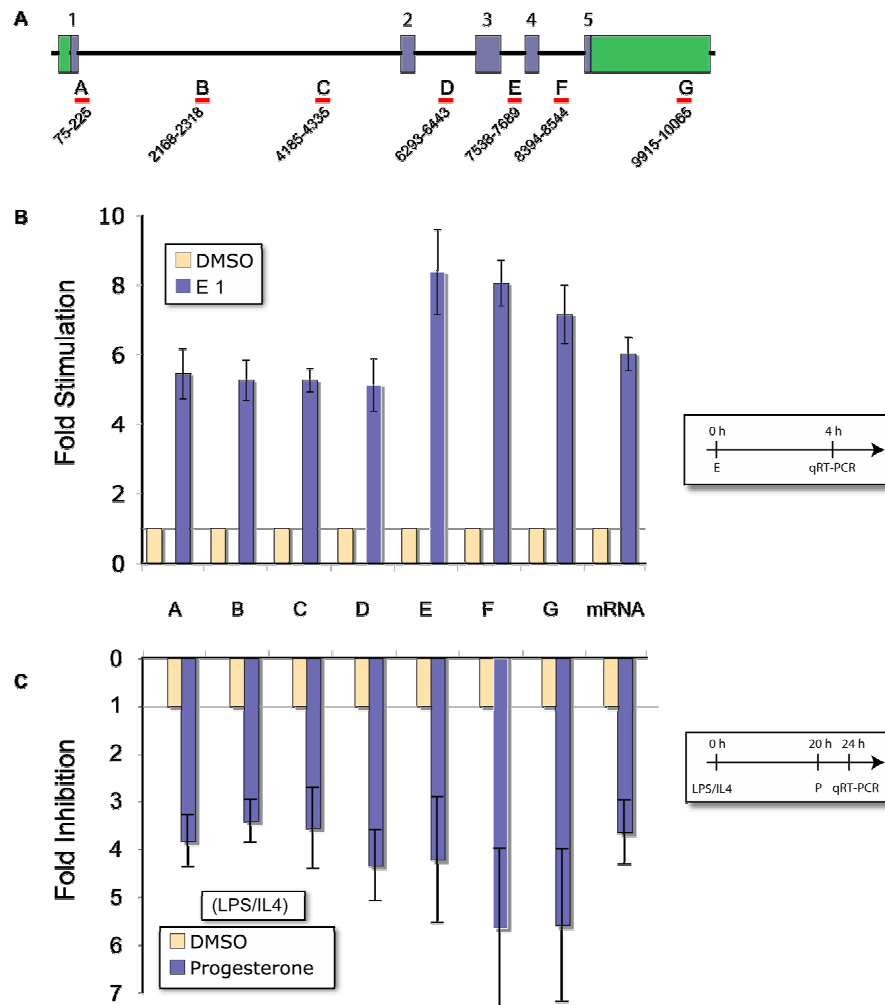
Appendix Figure 1. AID pre-mRNA processing in estrogen and progesterone treated cells.
(A) Schematic depiction of AID locus. Exons (blue) with corresponding numbers and UTRs (green) are indicated as boxes. Red lines with capital letters indicate the relative positions of the PCR products (A-G) along the AID locus. (B) The effects of estrogen on AID pre-mRNA. Unstimulated mouse splenic B-cells were treated with 1 nM (E 1) for 4 h, followed by analysis of AID pre-mRNA by qRT-PCR. Results are normalized to DMSO-treated cells for each PCR product. (B) The effects of progesterone on AID pre-mRNA. Mouse splenic B-cells were stimulated with LPS/IL4 as in Figure 11, followed by DMSO or progesterone treatment for 4 h. cDNA was synthesized, and analysed by qRT-PCR. Results are normalized to DMSO-treated cells for each PCR product. Data represents results from three independent experiments and error bars indicate standard deviations from the average. Time-lines of cell treatments are indicated beside the graph. Published in (Pauklin et al., 2009b) and (Pauklin and Petersen-Mahrt, 2009).
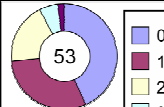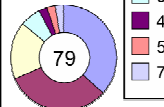
Appendix Figure 2.

Appendix Figure 2. Hormonal effects on AID mRNA and sIgM in Ramos. (A) The effect of estrogen on AID mRNA in Ramos HS13 cells. Cells were hormone depleted for 72 h prior to treatment with 1 nM estrogen (marked as E 1), 10 nM estrogen (E 10) or 100 nM estrogen (E 100) for 4 h. Data is represented as in Figure 11. Time-line of cell treatment is indicated beside the graph. Published in (Pauklin et al., 2009b). (B) The effects of estrogen on sIgM expression in Ramos. Sorted individual sIgM negative cells were grown in the presence of indicated amounts of estrogen for 3 weeks. Clones were analyzed for sIgM expression by flow cytometry. Each colored dot represents the relative proportion of sIgM positive cells per clone. (C) Inhibition of AID mRNA in RamosHS13 by progesterone treatment. Ramos cells were treated with 2 μM progesterone for 4 h and AID mRNA was analyzed by qRT-PCR. (D) Hormonal effects on sIgM expression in Ramos. Sorted individual sIgM negative cells were grown in the presence of indicated amounts of progesterone for 3 weeks. Clones were analyzed for sIgM expression by flow cytometry. Each colored dot represents the relative proportion of sIgM positive cells per clone. (E) The data from (D) expressed as fold inhibition on the median surface expression (DMSO set to 1). Cell growth analysis indicated that progesterone only had a marginal influence on cell replication, and results analysed on the number of cell duplications were not different (data not shown). NT - not treated. Published in (Pauklin and Petersen-Mahrt, 2009)
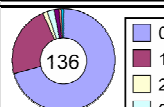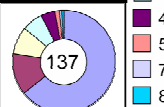
Appendix Figure 3.

A   VH Ramos

Treatment	Seq	BP	Mut	Mut/bp	% C:G	% Ts	Mutations/seq
DMSO	53	18468	50	2.7×10^{-3}	83.7	41.5	
Estrogen 100 nM	79	27360	100	3.7×10^{-3}	91.2	48.2	

B   CD95/Fas Ramos

Treatment	Seq	BP	Mut	Mut/bp
DMSO	78	54600	6	1.1×10^{-4}
Estrogen 100 nM	92	63700	12	1.9×10^{-4}

C   Sγ3 mouse splenic B cells
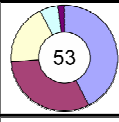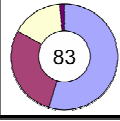
Treatment	Seq	BP	Mut	Mut/bp	% C:G	% Ts	Mutations/seq
DMSO	136	192694	73	3.8×10^{-4}	60.5	35.8	
Estrogen 10 nM	137	194051	126	$6.5 \times 10^{-4*}$	72.3	41.6	

$^{*}p < 0.02$

Appendix Figure 3. Estrogen increases the mutation frequency in Ig and non-Ig loci. (A and B) Ramos HS13 cell line was grown in the presence of indicated amounts of estrogen for 20 doublings, followed by cloning and sequencing of individual human VH (A) or human CD95/Fas (B) loci. The number of sequences analyzed (Seq), total base pairs (BP), number of mutations (Mut), mutation frequency per base pair (Mut/bp), overall percentage of mutations at C:G base pairs (% C:G), percentage of transitions (% Ts) and pie charts (mutations per sequence) are indicated for each of the treatments. As the overall mutations and percentage of transitions are already intrinsically very high in Ramos, estrogen treatment only modestly increases these percentages. (B) From single cell sorted clones in (A), genomic DNA was isolated and ~750 bp of the 5' CD95/Fas locus sequenced. The number of mutations in CD95/Fas locus in Ramos and does not allow for meaningful statistical analysis of mutations at C:G base pairs (% C:G) or percentage of transitions (% Ts). (C) Mouse splenic B-cells from AID+/- were isolated and stimulated with LPS and co-treated with DMSO or 10 nM estrogen for 72 h. Genomic DNA was amplified and 750 bp of the γ3 switch region sequenced. Table was generated as indicated in (A). The difference between the two samples 'mutation per base pair' was significant to p < 0.02 (two-tailed unpaired T-test). Published in (Pauklin et al., 2009b).

Appendix Figure 4.

A   VH Ramos

Treatment	Seq	BP	Mut	Mut/bp	% C:G	% Ts	Mutations/seq
DMSO	53	18468	50	2.7×10^{-3}	83.7	41.5	53 — 0, 1, 2, 3, 4
Progesterone 2 µM	83	28728	51	$1.8 \times 10^{-3*}$	72.0	36.1	83
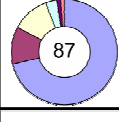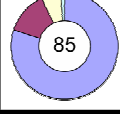
*p = 0.08

B   CD95/Fas Ramos

Treatment	Seq	BP	Mut	Mut/bp
DMSO	78	54600	6	1.1×10^{-4}
Progesterone 2 µM	92	64400	1	$0.16 \times 10^{-4*}$

*p = 0.05

C   Sγ3 (Balb/c)

Treatment	Seq	BP	Mut	Mut/bp	% C:G	% Ts	Mutations/seq
DMSO	87	118059	52	4.4×10^{-4}	68.6	41.2	87 — 0, 1, 2, 3, 4, 5
Progesterone 200 nM	85	115345	28	$2.4 \times 10^{-4*}$	57.1	25.0	85

*p = 0.04

Appendix Figure 4. The effect of progesterone on SMH in RAMOS HS13 cells and splenic B-cells. (A and B) Progesterone decreases the mutation frequency in the VH locus and the CD95/Fas locus. Ramos cells were grown in the presence of indicated amounts of progesterone for 3 weeks, followed by PCR amplification, cloning and sequencing of individual human VH (A) or human CD95/Fas (B) loci. The number of sequences analyzed (Seq), total base pairs (BP), number of mutations (Mut), mutation frequency per base pair (Mut/bp), overall percentage of mutations at C:G base pairs (% C:G), percentage of transitions (% Ts) and pie charts (mutations per sequence) are indicated for each treatment. Statistical significance was tested using a two-tailed unpaired T-test. (C) Progesterone inhibits SHM in the Sγ3 of mouse splenic cells. Mouse splenic cells were treated for 6 d with LPS and 200 nM progesterone. Genomic DNA was amplified and 750 bp of the Sγ3 region sequenced. Mutation frequencies are normalized to the control treatments with DMSO. The difference in mutation frequency in the Sγ3 loci of DMSO and progesterone treated spleen cells is significant, according to standard unpaired two-tailed T-test. Published in (Pauklin and Petersen-Mahrt, 2009).

Appendix Figure 5.
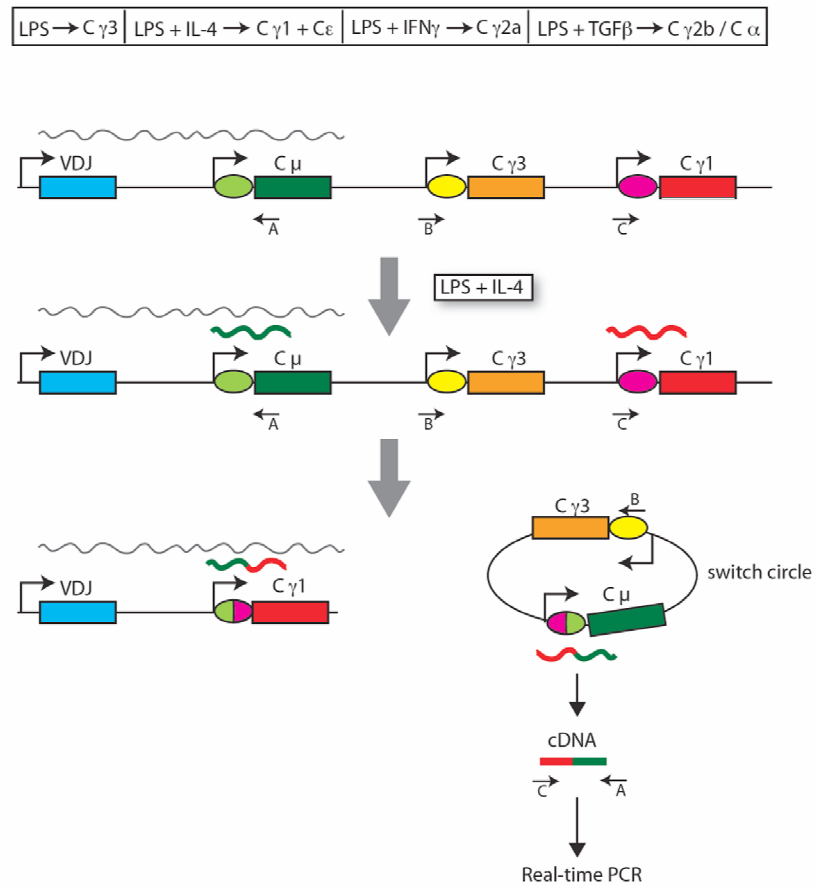
Appendix Figure 51 Schematic depiction of class switching. The formation of switch circles and production of 'circle' transcripts. The combinations of cytokines and LPS that lead to the switching of Ig isotypes are indicated in the top panel. Ovals indicate switch regions, arrows mark the promoters and rectangles are different constant region exons (Cμ, Cγ3, Cγ1). Class-switching can result in the formation switch circle, from which a hybrid transcript (switch circle transcript - wavy line) is expressed (Kinoshita et al., 2001). These transcripts can be reverse-transcribed for cDNA, and using specific primers (e.g. arrows marked with C and A for detecting the switching to Cγ1 isotype), quantitate switch circle formation by qRT-PCR. Published in (Pauklin et al., 2009b).

Appendix Figure 6.
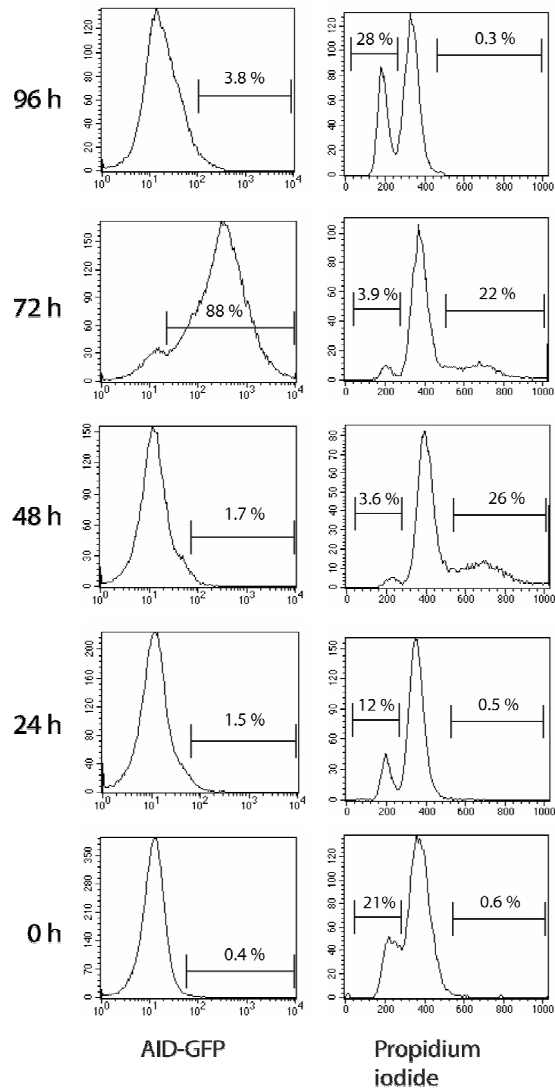
Appendix Figure 52 FACS analysis of nocodazole treatment. Yeast cells (see Figure 49) were analysed for eGFP-AID expression and DNA content with propidium-iodide by FACS analysis. Published in (Pauklin et al., 2009a).

Appendix Table 1.

Appendix Table 1. Identification of AID-interacting proteins. List of proteins identified by mass spectrometry in Figure 40. Cytoplasmic, nuclear and chromatin fractions were isolated from DT40 AID-FLAG-TEV-Myc DT40 cells, immunoprecipitated by anti-FLAG beads, visualized on a 4-12% gradient gel, cut out of the gel with a scalpel and analysed by mass spectrometry (see 3.2.4.6). Samples not treated with RNAse A prior to immunoprecipitation are marked with as (-)RNAse, samples treated with RNAse A are marked with (+)RNAse. Numbers in columons CL18- IP and AID IP mark the number of individual peptides identified for each protein in CL18- and AID-FLAG-TEV-Myc pulldown, respectively. Ribosomal proteins were not included in this table. Names in brackets and highlighted with Bold mark factors shown in the network of AID-interacting proteins (Figure 44).

Cytoplasm (-)RNAse

Protein ID	Accession number	Mw (kDa)	CL18-IP	AID IP
poly(A) binding protein, cytoplasmic 1 (Gallus gallus)	gi|71896197	71 kDa	1	11
Heterogeneous nuclear ribonucleoprotein K - Gallus gallus (Chicken)	HNRPK_CHICK	47 kDa	0	6
Guanine nucleotide-binding protein subunit beta-2-like 1 - Gallus gallus (Chicken) [RACK1]	GBLP_CHICK (+18)	35 kDa	1	4
Nucleophosmin - Gallus gallus (Chicken)	NPM_CHICK	33 kDa	1	4
14-3-3 protein beta/alpha - Gallus gallus (Chicken)	1433B_CHICK	28 kDa	0	2
Survival motor neuron protein - Gallus gallus (Chicken)	Q98SU9_CHICK	29 kDa	0	2
Cellular tumor antigen p53 - Gallus gallus (Chicken)	P53_CHICK	40 kDa	0	1
Importin subunit beta-1 - Homo sapiens (Human)	IMB1_HUMAN (+8)	97 kDa	0	1
Poly(rC)-binding protein 2 - Homo sapiens (Human)	PCBP2_HUMAN (+18)	39 kDa	0	1
GTP-binding nuclear protein Ran - Gallus gallus (Chicken)	RAN_CHICK (+128)	24 kDa	0	1
PREDICTED: similar to RAD54-like 2 (Macaca mulatta)	gi|109039362 (+138)	163 kDa	0	1
PREDICTED: similar to exportin 5 (Gallus gallus)	gi|118088066 (+5)	137 kDa	0	1

Nucleoplasm (-)RNAse

Protein ID	Accession number	Mw (kDa)	CL18-IP	AID IP
DEK oncogene (DNA binding) (Gallus gallus), hypothetical protein (Gallus gallus)	gi|60302770	42 kDa	0	9
Nuclease-sensitive element-binding protein 1 - Gallus gallus (Chicken)	YBOX1_CHICK (+2)	36 kDa	0	3
Lamin-B2 – Gallus gallus (Chicken)	LMNB2_CHICK	68 kDa	1	14
Nucleolin - Gallus gallus (Chicken)	NUCL_CHICK	76 kDa	0	10
Nuclear calmodulin-binding protein - Gallus gallus (Chicken)	Q9YHD2_CHICK	84 kDa	0	5
ATP-dependent RNA helicase DDX1 -	DDX1_CHICK	82 kDa	0	5

Gallus gallus (Chicken)				
RuvB-like 2 - Xenopus tropicalis (Western clawed frog) (Silurana tropicalis)	Q28GB7_XENTR	51 kDa	0	3
Cytoplasmic activation-proliferation-associated protein 1 - Gallus gallus (Chicken)	Q5XNV3_CHICK (+6)	78 kDa	0	2
Nuclear protein matrin 3 - Gallus gallus (Chicken)	Q8UWC5_CHICK	101 kDa	0	2
Pre-mRNA-processing-splicing factor 8 - Homo sapiens (Human)	PRP8_HUMAN (+22)	274 kDa	0	2
DNA topoisomerase I - Gallus gallus (Chicken)	P79994_CHICK	91 kDa	0	2
Cellular tumor antigen p53 - Gallus gallus (Chicken)	P53_CHICK	40 kDa	0	1
Importin subunit beta-1 - Homo sapiens (Human)	IMB1_HUMAN (+8)	97 kDa	0	1
Interleukin enhancer-binding factor 2 - Homo sapiens (Human)	ILF2_HUMAN (+16)	43 kDa	0	1
Splicing factor 3B subunit 4 - Homo sapiens (Human)	SF3B4_HUMAN (+10)	44 kDa	0	1
Histone H1.01 - Gallus gallus (Chicken)	H101_CHICK (+14)	22 kDa	0	1
PREDICTED: similar to endo/exonuclease Mre11 (Nasonia vitripennis)	gi|156554359	52 kDa	0	1
PREDICTED: similar to RAD54-like 2 (Macaca mulatta)	gi|109039362 (+138)	163 kDa	0	1
poly(A) binding protein, cytoplasmic 1 (Gallus gallus)	gi|71896197	71 kDa	0	1

Chromatin (-)RNAse

Lamin-B2 – Gallus gallus (Chicken)	LMNB2_CHICK	68 kDa	0	4
poly(A) binding protein, cytoplasmic 1 (Gallus gallus)	gi|71896197	71 kDa	0	3
Nucleophosmin - Gallus gallus (Chicken)	NPM_CHICK	33 kDa	0	2
Nucleolin - Gallus gallus (Chicken)	NUCL_CHICK	76 kDa	0	2
RuvB-like 2 - Xenopus tropicalis (Western clawed frog) (Silurana tropicalis)	Q28GB7_XENTR	51 kDa	0	1
Splicing factor 3a subunit 2 - Gallus gallus (Chicken)	Q66VY4_CHICK (+11)	35 kDa	0	1
Splicing factor 3B subunit 2 - Homo sapiens (Human)	SF3B2_HUMAN (+4)	98 kDa	0	1
Nuclear calmodulin-binding protein - Gallus gallus (Chicken)	Q9YHD2_CHICK	84 kDa	0	1
Cytoplasmic activation-proliferation-associated protein 1 - Gallus gallus (Chicken)	Q5XNV3_CHICK (+6)	78 kDa	0	1
Nuclear protein matrin 3 - Gallus gallus (Chicken)	Q8UWC5_CHICK	101 kDa	0	1
Heterogeneous nuclear ribonucleoprotein A3 - Homo sapiens (Human)	ROA3_HUMAN (+16)	40 kDa	0	1

DNA-directed RNA polymerase II subunit RPB3 - Bos taurus (Bovine)	RPB3_BOVIN (+6)	31 kDa	0	1
Spliceosome RNA helicase BAT1 - Gallus gallus (Chicken)	UAP56_CHICK (+63)	49 kDa	0	1
FACT complex subunit SPT16 (Facilitates chromatin transcription complex subunit spt16)	gi|110287969 (+21)	118 kDa	0	1
DNA excision repair protein Rad16, putative	gi|119482550	111 kDa	0	1

Nucleoplasm (+)RNAse treatment

matrin 3 (Gallus gallus), nuclear protein matrin 3 (Gallus gallus)	gi|45383822	101 kDa	1	9
PREDICTED: interleukin enhancer binding factor 3 isoform 7 (Pan troglodytes)	gi|114675397 (+18)	96 kDa	0	3
nuclear calmodulin-binding protein (Gallus gallus)	gi|3822553	84 kDa	1	4
DEK oncogene (DNA binding) (Gallus gallus), hypothetical protein (Gallus gallus)	gi|60302770	42 kDa	0	4
Splicing factor 3B subunit 3 - Bos taurus (Bovine)	SF3B3_BOVIN (+4)	136 kDa	2	9
Splicing factor 3B subunit 1 - Homo sapiens (Human)	SF3B1_HUMAN (+2)	146 kDa	0	4
Nuclease-sensitive element-binding protein 1 - Gallus gallus (Chicken)	YBOX1_CHICK (+28)	36 kDa	0	1
Nuclear protein matrin 3 - Gallus gallus (Chicken)	Q8UWC5_CHICK	101 kDa	0	6
Nuclear cap-binding protein subunit 1 - Gallus gallus (Chicken)	NCBP1_CHICK	93 kDa	0	3
Pre-mRNA-processing factor 6 - Homo sapiens (Human)	PRP6_HUMAN (+12)	107 kDa	0	1
NIK- and IKBKB-binding protein - Bos taurus (Bovine)	NIBP_BOVIN (+6)	127 kDa	0	1
Putative uncharacterized protein - Gallus gallus (Chicken)	Q5ZIR3_CHICK (+11)	155 kDa	0	1
Nuclear calmodulin-binding protein - Gallus gallus (Chicken)	Q9YHD2_CHICK	84 kDa	0	1
Transport protein particle subunit TMEM1 - Homo sapiens (Human)	TMEM1_HUMAN (+13)	142 kDa	0	1

Chromatin (+)RNAse treatment

RNA polymerase II-associated factor 1 homolog - Bos taurus (Bovine) **[PAF1]**	PAF1_BOVIN (+5)	60 kDa	0	4
DNA-directed RNA polymerase II subunit RPB1 - Homo sapiens (Human) **[POLR2A]**	RPB1_HUMAN (+6)	217 kDa	0	1
DNA-directed RNA polymerase II subunit RPB2 - Homo sapiens (Human)	RPB2_HUMAN (+5)	134 kDa	0	5

Protein	Accession	MW		
[POLR2B]				
Pre-mRNA-processing factor 19 – Gallus gallus (Chicken) **[PRPF19]**	PRP19_CHICK	55 kDa	0	6
Pre-mRNA-processing factor 6 - Homo sapiens (Human) **[PRPF6]**	PRP6_HUMAN (+12)	107 kDa	0	1
Lamin-B1 - Gallus gallus (Chicken) **[LMNB1]**	LMNB1_CHICK	67 kDa	0	1
Transcription elongation factor SPT5 - Gallus gallus (Chicken) **[SUPT5H]**	SPT5H_CHICK	120 kDa	0	11
Transcription elongation factor SPT6 - Homo sapiens (Human) **[SUPT6H]**	SPT6H_HUMAN (+2)	199 kDa	0	4
Transcription elongation factor SPT6 - Homo sapiens (Human) **[SUPT6H]**	SPT6H_HUMAN (+2)	199 kDa	0	4
DNA topoisomerase I - Gallus gallus (Chicken) **[TOP1]**	P79994_CHICK	91 kDa	0	2
FACT complex subunit SPT16 - Homo sapiens (Human) **[SUPT16H]**	SP16H_HUMAN (+6)	120 kDa	0	1
FACT complex subunit SSRP1 - Gallus gallus (Chicken) **[SSRP1]**	SSRP1_CHICK (+13)	80 kDa	0	1
ATP-dependent RNA helicase DDX1 - Gallus gallus (Chicken) **[DDX1]**	DDX1_CHICK (+15)	82 kDa	0	1
Nucleolin - Gallus gallus (Chicken) **[NCL]**	NUCL_CHICK	76 kDa	3	9
PREDICTED: interleukin enhancer binding factor 3 isoform 7 (Pan troglodytes)	gi|114675397 (+18)	96 kDa	0	4
nuclear calmodulin-binding protein (Gallus gallus)	gi|3822553	84 kDa	0	0
DEK oncogene (DNA binding) (Gallus gallus), hypothetical protein (Gallus gallus)	gi|60302770	42 kDa	1	2
Highly divergent homeobox - Gallus gallus (Chicken)	HDX_CHICK	77 kDa	0	3
U5 small nuclear ribonucleoprotein 200 kDa helicase - Homo sapiens (Human)	U520_HUMAN (+1)	245 kDa	1	7
Elongation factor 2 - Gallus gallus (Chicken)	EF2_CHICK	95 kDa	0	6
Splicing factor 3B subunit 1 - Homo sapiens (Human)	SF3B1_HUMAN (+2)	146 kDa	1	3
Nuclease-sensitive element-binding protein 1 - Gallus gallus (Chicken)	YBOX1_CHICK (+28)	36 kDa	0	1
RNA polymerase II-associated factor 1 homolog - Bos taurus (Bovine)	PAF1_BOVIN (+5)	60 kDa	0	4
Lamin-B2 – Gallus gallus (Chicken)	LMNB2_CHICK	68 kDa	0	4
Pre-mRNA-processing factor 6 - Homo sapiens (Human)	PRP6_HUMAN (+12)	107 kDa	0	1
Lamin-B1 - Gallus gallus (Chicken)	LMNB1_CHICK	67 kDa	0	1