Ictalurus Punctatus Activation-Induced Cytidine Deaminase Expression

by

Holly L. Saunders

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

In

Physiology, Cell, and Developmental Biology

Department of Biological Sciences

University of Alberta

Abstract

Antigen-specific recognition and memory, the hallmarks of conventional mammalian adaptive immunity, are found in all jawed vertebrates along with all elements central to cognate adaptive immune responses (e.g. B-cells). During T-cell-dependent humoral immune responses in mammals, antigen-specific recognition is significantly enhanced by antibody affinity maturation. The mammalian antibody affinity maturation paradigm, derived from mouse/human, involves genetic modification of the immunoglobulin (Ig) genes by somatic hypermutation (SHM) and class switch recombination (CSR) coupled with phenotypic selection for B-cells that express higher-affinity antibody. Antigen-activated B-cells undergo these processes while sequestered in germinal centers, which are transient, specialized microenvironments that develop in secondary lymphoid tissues. These histologically obvious sites also harbour cognate T-cells and follicular dendritic cells in an organization that is believed to be essential to efficient selection and, consequently, to substantial increases in antibody affinity.

Although SHM and CSR are discrete events, they share several features including an indispensable requirement for the mutator enzyme, activation-induced cytidine deaminase (AID). In germinal center B-cells, the AID enzyme deaminates cytosine to uracil in single-stranded (ssDNA) of the transcribed Ig genes. This pre-mutagenic uracil DNA lesion is a target of ubiquitously expressed DNA damage tolerance and repair molecules, which, ultimately, achieve SHM and CSR outcomes.

Information about the immune response in lower vertebrates (sharks, bony fishes, amphibians) is relatively sparse, but they appear to exhibit limited if any antibody affinity maturation. Although conventional germinal centers have not been detected in lower vertebrates, SHM has been detected in sharks and both SHM and CSR have been detected in amphibians. In the teleost bony fishes, CSR cannot occur due to the absence of additional constant regions in the Ig heavy chain (IgH) gene, but their SHM status is unknown.

To better understand antibody affinity maturation in teleosts, the following questions need to be answered: Are teleosts able to somatically hypermutate their Ig genes? And, if so, how do teleosts balance the risks and benefits (i.e. prevent autoimmunity versus improve antibody affinity) of this highly mutagenic process? To address these questions, we sought an AID homologue in the channel catfish (*Ictalurus punctatus*), a representative teleost. Since AID is the only B-cell-specific factor required for SHM in mammals, the identification of an AID homologue in catfish could be used to determine if teleosts are capable of SHM. It could also be used as a marker to specifically identify *in situ* the teleost equivalents of mammalian germinal center B-cells as indicators of sites of possible germinal center analogues.

We found two lines of evidence that support the proposal that catfish undergo AID-mediated SHM. First, catfish express an AID homologue in a tissue expression pattern that is consistent with AID-mediated mutation playing a role in immune function. The catfish AID homologue has 57% amino acid identity and 81% amino acid similarity with mouse/human AIDs and maintains strong

conservation of amino acids that have been determined to be important for cytidine deaminase, SHM, and CSR functions. After native AID gene expression had been upregulated in a catfish immortal B-cell line, DNA sequence comparison of the endogenous IgH gene variable region exon in subclones revealed the presence of SHM with the mutation signature of AID. Thus, catfish AID is catalytically active *in vivo* and, as in mammals, is targeted to Ig genes.

Second, cell aggregates observed in catfish spleen sections appear to be broadly reminiscent of mammalian germinal centers. Distinct cells that express AID mRNA colocalize in clusters with melanomacrophage cells, which trap antigen and are putative teleost analogues of mammalian follicular dendritic cells. That B-cells undergo AID-mediated SHM at sites where antigen is “trapped” suggests the presence of concomitant phenotypic selection. However, the intense selection that occurs in mammalian germinal centers is accompanied by crucial features, such as extensive B-cell apoptosis, that appear to be absent in catfish. Notably, in mouse, blocking B-cell apoptosis in germinal centers severely diminishes antibody affinity maturation.

In conclusion, we have shown that catfish diversify their Ig genes via AID-dependent SHM and that cells that express AID reside in distinct cellular aggregations partially reminiscent of mammalian germinal centers. Since catfish can somatically modify their Ig genes, the apparently lesser antibody affinity maturation of teleosts may result from relatively inefficient selection for higher-affinity variants.

Preface

This thesis is an original work by Holly L. Saunders. Research ethics approval for the research projects that comprise this thesis was obtained from the University of Alberta Animal Care and Use Committee under the project names “Immunoglobulin Affinity Maturation in Fishes” (2003 to 2006) and “Mechanisms of Humoral Immunity in Fishes” (2006 to 2008), Nos. 430312 through 430812. Live specimens were maintained by the University of Alberta Department of Biological Sciences Aquatic Facility staff according to guidelines set forth by the Canadian Council on Animal Care.

Some of the data and materials obtained from the research projects that comprise this thesis have contributed to articles published in refereed journals.

Article 1

Cloning and expression of the AID gene in the channel catfish. Saunders HL, Magor BG. Developmental and Comparative Immunology, 2004, 28(7-8): 657-663. I cloned the activation-induced cytidine deaminase (AID) homologue from the cDNA of the channel catfish. I measured the relative expression of mRNA from the AID homologue gene in various channel catfish tissues (See Results and Discussion).

Article 2

Evolution of class switch recombination function in fish activation-induced cytidine deaminase, AID. Wakae K, Magor BG, Saunders H, Nagaoka H, Kawamura A, Kinoshita K, Honjo T, Muramatsu M. International Immunology, 2006, 18(1): 41-47. I provided a plasmid bearing the cloned cDNA of the AID homologue gene from channel catfish for the research detailed in this article. This article reports the relative SHM and CSR activities of AID homologues from various organisms (mouse, human, chicken, zebrafish, catfish) when expressed in a mammalian (i.e. mouse) cellular context. As part of this research collaboration, I performed complementary experiments in a channel catfish

immortal B-cell line to examine the functional capacity of AID homologues from various organisms (channel catfish, zebrafish, human) when expressed in a teleost cellular context (unpublished; See Results, Discussion, and Appendix).

Article 3

The cellular context of AID-expressing cells in fish lymphoid tissues. Saunders HL, Oko AL, Scott AN, Fan CW, Magor BG. Developmental and Comparative Immunology, 2010, 34(6): 669-676. I prepared and administered vaccinations to the experimental specimens prior to harvest of their tissues and cells. All parties contributed to specimen dissection and harvest of tissues and cells for preservation and analysis. I performed *in situ* hybridization in channel catfish tissues. I exposed channel catfish immortal B-cell lines to various mitogens and assessed their ability to upregulate AID mRNA expression (See Results and Discussion).

Article 4

Differences in the enzymatic efficiency of human and bony fish AID are mediated by a single residue in the C terminus modulating single-stranded DNA binding. Dancyger AM, King JJ, Quinlan MJ, Fifield H, Tucker S, Saunders HL, Berru M, Magor BG, Martin A, Larijani M. Journal of the Federation of American Societies for Experimental Biology, 2012, 26(4): 1517-1525. I used mitogens to upregulate native AID expression in a channel catfish immortal B-cell line. I then sequenced the variable region of the endogenous IgH gene and analyzed it for mutations (See Results and Discussion).

Dedication

To all the fishies who gave their lives in furtherance of science and to my old folks, who gave their money in furtherance of my education.

Acknowledgements

Foremost, I want to thank my graduate supervisor, Dr. Brad Magor, for sharing with me his accumulated wisdom and his love of science.

Furthermore, I want to thank: the members of my graduate committee, Dr. Linda Pilarski, Dr. Linda Reha-Krantz, Dr. Robert (Jack) von Borstel, Dr. Hanne Ostergaard, and Dr. David Evans, for their time, encouragement, and guidance; the examiners and chairs of my PhD Candidacy and Defense exams, Dr. Kristi Baker, Dr. Mike Criscitiello, Dr. Patrick Hanington, Dr. Kevin Kane, Dr. James Stafford, and Dr. Keith Tierney, for their time and helpful feedback; Dr. Miodrag (Mike) Belosevic, Dr. Mario Feldman, and Dr. Kathy Magor for their kindness in allowing me the use of their laboratories, equipment, and expertise; the many helpful members of the Magor lab whose own research augmented my thesis research, Gregory Eustace, Charity Fan, Andrea Oko, Amanda Scott, Daniela Villota-Herdioza, Angela Zolner, and Jeffery Zorn; my collaborators; and, the multitude of my peers, support staff, and faculty who provided technical, administrative, and academic assistance.

I was very fortunate to have been the recipient of several scholarships and bursaries for which I want to express my inestimable gratitude to the agencies whence they were bestowed: Natural Sciences and Engineering Research Council (NSERC) of Canada, University of Alberta Faculties of Science and Graduate Studies and Research, University of Alberta Graduate Student Association, International Society for Developmental and Comparative Immunology, and Canadian Society for Immunology.

Finally, I am very grateful to Dr. Dion Durnford, Dr. Kevin Englehart, Dr. Drew Rendall, and others at the University of New Brunswick and the University of Alberta who went out of their way during the 2020 Coronavirus pandemic to arrange for me to defend my thesis by videoconference from UNBF.

“So long, and thanks for all the fish.” - Douglas Adams

Table Of Contents

List Of Tables

List Of Figures

List Of Abbreviations

Abbreviation	Definition
αIgM	Anti-IgM antibody
5-mC	5-Methylcytosine
Ab	Antibody
aicda	Activation-Induced Cytidine Deaminase gene
AID	Activation-Induced Cytidine Deaminase
AIM V	Adoptive Immunotherapy Media-V
APE	Apurinic/Apyrimidinic Endonuclease
apoB48	Apolipoprotein B48
apoB100	Apolipoprotein B100
APOBEC1	Apolipoprotein B mRNA-Editing Catalytic Subunit 1
BCIP	5-Bromo-4-Chloro-3-Indolyl Phosphate
Bcl-2	B-Cell Lymphoma 2
Bcl-6	B-Cell Lymphoma 6
Bcl-xL	B-Cell Lymphoma-Extra Large
BCR	B-Cell Receptor
BER	Base Excision Repair
BOB.1	B-Cell-Specific Octamer-Binding Protein-1 (OCA-B, OBF-1)[1]
BSA	Bovine Serum Albumin
ccm	Cell-Conditioned Medium
CD4	Cluster of Differentiation 4
CD19	Cluster of Differentiation 19

CD40	Cluster of Differentiation 40
CD79α	Cluster of Differentiation CD79α (Ig Alpha)
CD79β	Cluster of Differentiation CD79β (Ig Beta)
CD154	Cluster of Differentiation 154 (CD40 Ligand)
CDR	Complementarity Determining Region
CDS	Coding Sequence
Cf	Catfish
cfm	Complete Catfish Medium
CI	Calcium Ionophore
ConA	Concanavalin A
CR2	Complement Receptor 2 (Cluster of Differentiation 21)
CSR	Class Switch Recombination
Ctrunc	Carboxy-Terminal Truncated
CXCL12	C-X-C Motif Chemokine Ligand 12
CXCL13	C-X-C Motif Chemokine Ligand 13
CXCR4	C-X-C Motif Chemokine Receptor 4
CXCR5	C-X-C Motif Chemokine Receptor 5
DEPC-treated mQH2O	mQH2O treated with diethyl pyrocarbonate to inactivate RNases, then autoclaved to sterilize it and eliminate residual diethyl pyrocarbonate.
DIG	Digoxigenin
DIG-11-UTP	Digoxigenin-11-uridine-5'-triphosphate
dUTP	2’-Deoxyuridine 5’-Triphosphate
dUTPase	Deoxyuridine Triphosphatase

EDTA	Ethylenediaminetetraacetic Acid
Ei	Intronic Enhancer of IgH gene (Eμ in mouse)
FACS	Fluorescence-Activated Cell Sorter
FBS	Fetal Bovine Serum
FCA	Freund's Complete Adjuvant
FDC	Follicular Dendritic Cell
FFPE	Formalin-Fixed Paraffin-Embedded
FIA	Freund's Incomplete Adjuvant
FITC-BSA	Fluorescein Isothiocyanate conjugated to Bovine Serum Albumin
FR	Framework Region
HE	Haematoxylin and Eosin Stains
HIGM	Hyper-Immunoglobulin M Syndrome — heterogeneous etiology with sub-types, due to defects in specific genes, denoted by a numeric suffix. Type 2 is due to AID defects.
Ig	Immunoglobulin
IgA	Immunoglobulin A
IgD	Immunoglobulin D
IgE	Immunoglobulin E
IgG	Immunoglobulin G
IgH	Immunoglobulin Heavy Chain
IgM	Immunoglobulin M
IgNAR	Immunoglobulin New Antigen Receptor
IgY	Immunoglobulin Y
Ip	*Ictalurus punctatus* (Channel Catfish)

IpCD4L1	*Ictalurus punctatus* CD4-Like 1
IpCD4L2	*Ictalurus punctatus* CD4-Like 2
ISH	*In Situ* Hybridization
IVT	*In Vitro* Transcription
L-15	Leibovitz's culture medium. This formulation contains the standard fifteen levorotatory amino acids plus L-glutamine.
LTα	Lymphotoxin-alpha
MALT	Mucosa-Associated Lymphoid Tissue
MAR	Matrix Attachment Region
MHC II	Major Histocompatibility Complex Class II
MMR	Mismatch Repair
mQH2O	Water obtained from Millipore Milli-Q water purification device and autoclaved to sterilize it.
MSH2	MutS Homologue 2
MSH6	MutS Homologue 6
Mug	Mismatch-Specific Uracil-DNA Glycosylase
MutS	Mutator S
NBF	Neutral Buffered Formalin
NBT	Nitroblue Tetrazolium
NCBI	National Center for Biotechnology Information
ncMMR	Non-Canonical Mismatch Repair
NES	Nuclear Export Signal
NHEJ	Non-Homologous End-Joining
NLS	Nuclear Localization Signal

OBF-1	Oct-Binding Factor-1 (BOB.1, OCA-B)
OCA-B	Oct Coactivator from B-cells (BOB.1, OBF-1)
PBL	Peripheral Blood Lymphocyte
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline + 0.1% Tween 20
Pfu	*Pyrococcus furiosus* DNA polymerase
PKA	Protein Kinase A
PMA	Phorbol 12-Myristate 13-Acetate
RAG	Recombination-Activating Gene
RACE	Rapid Amplification of cDNA Ends
RAD51B	Radiation-Sensitive 51B
Rev1	UV-Reversionless 1
RPA	Replication Protein A
RSS	Recombination Signal Sequence
RT	Reverse Transcription
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
R/S	Replacement/Silent Mutation Ratio
SDS	Sodium Dodecyl Sulfate
SHM	Somatic Hypermutation
SMUG1	Single-Strand-Selective Monofunctional Uracil-DNA-Glycosylase
SSC	Sodium Chloride Sodium Citrate Buffer
SSPE	Sodium Chloride Sodium Phosphate EDTA Buffer
Taq	*Thermus aquaticus* DNA polymerase
tBLASTx	Translated Basic Local Alignment Search Tool X. In this BLAST

	variant, “X” indicates that both the query sequence and database have been translated into the six possible reading frames.
TCR	T-Cell Receptor
TCRα	T-Cell Receptor Alpha
TCRβ	T-Cell Receptor Beta
TD-PCR	Touchdown-Polymerase Chain Reaction
TM	Transmembrane
TMS	Tricaine Methanesulfonate
TNFα	Tumor Necrosis Factor Alpha
TNFRI	Tumor Necrosis Factor Receptor 1
TRE	Tetracycline Response Element
Tris	Tris(hydroxymethyl)aminomethane
tTA	Tetracycline-Responsive Transcriptional Activator
TUNEL	Terminal Deoxynucleotidyl Transferase-Mediated dUTP Nick-End Labeling
Ugi	Uracil-DNA Glycosylase Inhibitor
UNG	Uracil-DNA Glycosylase
UNG2	Uracil-DNA Glycosylase 2
UTR	Untranslated Region
VDJ	Variable Diversity Joining (recombined IgH variable region)
WGS	Whole Genome Sequence
XAID	Shorthand term to encompass all of the different AID sequences used herein, where “X” represents any of the different organisms from which the AID sequence has been derived. These are full-length cDNA sequences unless the carboxy-terminal truncated

form is indicated by inclusion of “Ctrunc”.

XRCC2 X-Ray Repair Cross Complementing 2

XRCC3 X-Ray Repair Cross Complementing 3

[1] Parentheses indicate aliases of molecules.

1. Introduction

1.1 Antibody Affinity Maturation Events Have Not Been Defined In Teleosts

Antigen-specific recognition and memory, the hallmarks of conventional mammalian adaptive immunity, are found in all jawed vertebrates along with all elements central to cognate adaptive immune responses [Clem, 1990; Diaz, 1998]. This includes B-cells and the immunoglobulins (Ig) they produce in the form of membrane-bound B-cell receptor (BCR) and secreted antibody (Ab). Although adaptive immunity in non-mammalian jawed vertebrates appears to largely recapitulate that of mammals, there are some notable differences [Clem, 1990; Iwama, 1996]. For example, the T-cell-dependent humoral immune responses of lower vertebrates (sharks, bony fishes, amphibians) appear to generate relatively limited if any antibody affinity maturation [Cain, 2002; Diaz, 1998; Hsu, 1998; Iwama, 1996; Kaattari, 2002].

In mammals, the significant enhancement of antigen-specific recognition through antibody affinity maturation is an essential component of an efficient response [Tarlinton, 2000; Wabl, 1999]. In the mouse/human paradigm, antibody affinity maturation relies on genetic modification of the Ig genes by somatic hypermutation (SHM) and class switch recombination (CSR) coupled with phenotypic selection for B-cells that express higher-affinity antibody [Tarlinton, 2000; Wabl, 1999]. Activation-induced cytidine deaminase (AID), an enzyme that deaminates cytosine (C) to uracil (U) in single-stranded DNA (ssDNA), initiates both Ig gene modifications [Muramatsu, 2000; Muramatsu, 1999]. Replication or resolution of the pre-mutagenic U DNA lesions by different replicative DNA polymerases or DNA damage tolerance and repair mechanisms generates SHM in the Ig variable region exon and CSR of the constant region exons [Petersen-Mahrt, 2002]. Phenotypic selection for B-cells that express higher-affinity antibody is facilitated by germinal centers, which are transient, histologically obvious sites that develop in secondary lymphoid tissues [Kelsoe, 1996; Tarlinton, 1998; Tarlinton, 2000; Wabl, 1999]. Germinal centers contribute to

antibody affinity maturation by sequestering hypermutating B-cells in proximity to cognate T-cells and follicular dendritic cells (FDC), which provide the B-cells with regulatory signals and access to concentrated antigen in the form of immune-complexes [Kelsoe, 1996; Park, 2005; Tarlinton, 1998; Tarlinton, 2000; Wabl, 1999].

In mammals, defects that disrupt Ig gene modification or selection can reduce antibody affinity maturation [Diaz, 2002; Durandy, 2005; Matsumoto, 1997; Meffre, 2001]. This suggests that the relatively limited antibody affinity maturation detected in lower vertebrates may arise from differences in their ability to generate and/or select mutants. Intriguingly, although SHM has been detected in sharks and both SHM and CSR have been detected in amphibians, conventional germinal centers have not been observed in lower vertebrates [Diaz, 1998; Diaz, 2001; Hsu, 1998; Iwama, 1996]. The occurrence of Ig gene modifications in these organisms hints that some form of selection likely exists to balance the risks and benefits of hypermutation (i.e. prevent autoimmunity versus improve antibody affinity). But the inability to detect germinal centers by conventional means indicates that the selection mechanism may not be as efficient as that of mammals.

In the teleost bony fishes, CSR cannot occur due to the lack of additional constant region exons in the Ig heavy chain (IgH) gene, but whether they have SHM is unknown [Cain, 2002; Diaz, 1998; Hsu, 1998; Iwama, 1996; Magor, 1999]. To determine whether teleosts have SHM and a germinal center analogue, we sought an AID homologue in the channel catfish (*Ictalurus punctatus*), a representative teleost. Since AID is the only B-cell-specific factor required for SHM in mammals, the identification of an AID homologue in catfish could be used to determine if teleosts are capable of SHM. It could also be used as a marker to specifically identify *in situ* the teleost equivalents of mammalian germinal center B-cells as indicators of sites of possible germinal center analogues.

1.2 The Mammalian Humoral Immune System: The Benchmark For Comparative Research In Lower Vertebrates

The main components of the conventional adaptive immune system, including those of humoral immunity, are found in all extant classes of gnathostomes (i.e. jawed vertebrates) [Clem, 1990; Diaz, 1998]. However, because mammals are a primary focus of medical research, most aspects of the adaptive immune response, including the T-cell-dependent humoral immune response, are currently best understood in mammals. This is also true for important complementary areas of research (e.g. genome sequence, endocrine function, and organ systems) that have a bearing on the development and function of the adaptive immune system. Thus, the mammalian adaptive immune system is the only benchmark available to use to evaluate the adaptive immune systems of lower vertebrates (e.g. the channel catfish examined herein as a teleost model organism [Clem, 1990; Clem, 1996]) and it is common for the assumption to be made that the systems of lower vertebrates approximate those of mammals until the receipt of empirical data from lower vertebrates [Zwollo, 2005].

Accordingly, the Introduction is divided into two main sections. First, the mammalian paradigm of T-cell-dependent humoral immunity is presented. This has been derived primarily from studies of the laboratory mouse and human and many of the details have yet to be fully elucidated. Also noted are relevant deviations from the mouse/human paradigm among other mammals and another higher vertebrate, the chicken [Arakawa, 2002; Diaz, 2002; Neuberger, 2003; Neuberger, 1995]. This review has been customized to highlight those elements and events germane to an exploration of and comparison with potential AID-mediated antibody affinity maturation in channel catfish. Second, a review of the data from teleosts, with an emphasis on channel catfish, and other lower vertebrates is presented. Then the research proposal is outlined.

Although facilitated by T-cells, the mammalian paradigm of the T-cell-dependent humoral immune response centers on B-cells, which produce the actual humoral component of the response (i.e. antibody) [Manser, 2004;

Tarlinton, 2000]. Each B-cell expresses immunoglobulin — in the form of membrane-bound BCR or secreted antibody — with a unique antigen-recognition domain [Bassing, 2002]. Thus, the B-cell population and the resultant pool of serum antibodies are highly diverse. Immunoglobulin is both an essential detector of non-self foreign molecules (i.e. antigens) [Besmer, 2004], which typically connote the presence of pathogens (i.e. infection), and a principal conductor of the ensuing immune response. Detection of cognate antigen through binding to BCR initiates responses that can lead to B-cell activation [Manser, 2004; Tarlinton, 1998; Tarlinton, 2000]. Activation of B-cells is a preliminary step toward complex AID-mediated antibody affinity maturation processes whereby the antigen-recognition and isotype-determining effector domains of the antibody will be genetically altered to better fit the specific antigen to both improve antigen binding and refine the immune response [Green, 1998; Hsu, 2006; Neuberger, 2003]. Significant antibody affinity maturation is contingent on AID-mediated genetic alteration of the Ig genes being accompanied by robust antigen-dependent selection of B-cells [Tarlinton, 2000; Wabl, 1999].

The maturation of antibody affinity in response to antigen exposure enhances not only the current immune response — by developing a population of plasma B-cells that each express distinct antibody with improved capacity to serve antibody's myriad functions [Pham, 2005], such as opsonization, neutralization, complement activation, and antibody-dependent cell-mediated cytotoxicity — but also any future immune response that develops upon subsequent re-exposure to the same antigen — by developing an augmented population of long-lived memory B-cells that each express distinct BCR with improved sensitivity to the antigen.

1.2.1 Mammalian Primary Antibody Repertoire Generation

An antibody — composed of two identical immunoglobulin heavy chains (IgH) and two identical immunoglobulin light chains (IgL), the polypeptide products of

the IgH and IgL genes, respectively — can be divided into two functional domains that manage either antigen recognition or immune response (i.e. effector function), correspondingly termed the variable and constant domains (Figure 1) [Hsu, 2006]. This division of functions in the antibody molecule is reflected in the coding DNA sequence (CDS) of the IgH gene wherein the variable region exon encodes the portion of the immunoglobulin protein that contributes to the formation of the antibody variable domain and the constant region exons encode the portion of the immunoglobulin protein that forms the constant domain of the antibody molecule (Figure 1) [Hsu, 2006]. Additionally, alternative splicing of the IgH pre-mRNA transcript determines whether the translated IgH protein contributes to the formation of membrane-bound BCR or secreted antibody.

The amino-terminal variable domain of the antibody is responsible for antigen recognition by means of two antigen-binding sites that are identical in antibodies produced by an individual B-cell but can be exceedingly diverse among antibodies produced from different B-cells. This diversity, a crucial element of adaptive immunity, is the product of the genetic mechanisms (described below) that create distinct variable region exons in the IgH and IgL genes within each B-cell during its development and the contribution that each unique IgH and IgL polypeptide makes when they pair to form the antigen-binding site. Additional variable region alteration (i.e. SHM) can occur during antibody affinity maturation after the mature B-cell encounters its cognate antigen (Figure 2) [Besmer, 2004; Hsu, 2006].

The variable domain can be sub-divided further into framework regions (FR), which are primarily important for the preservation of antibody structural integrity, and complementarity determining regions (CDR), which are largely important for the interaction of antibody with discrete sites on the antigen known as epitopes (i.e. antigen-binding specificity) [Clark, 2006]. It is approximately the strength of this antibody-antigen interaction that dictates the affinity of an antibody's binding-site for its cognate antigen. However, the overall antigen-binding capabilities of any given antibody may be enhanced by their inherent possession of not one but

two antigen-binding sites, a phenomenon known as functional affinity or avidity [Iwama, 1996; Torres, 2008].

The carboxy-terminal constant domain of the antibody, which determines the antibody isotype or class, is primarily responsible for effector functions (e.g. complement fixation) [Barreto, 2003; Hsu, 2006; Poltoratsky, 2000; Tashiro, 2001] that better target the immune response to the particular type of pathogen that initiated it. Studies (reviewed in [Torres, 2008]) have shown, however, that the constant domain can exert an effect directly on antigen-binding affinity in the variable domain. The antibody constant domain, comprised solely of IgH, exhibits diversity but is not intrinsically diverse like the antibody variable domain for which the encoding exon is assembled afresh in each developing B-cell from sub-exons located upstream of the gene (described below) [Hsu, 2006]. Instead, after the mature B-cell encounters its cognate antigen, the exons of the germline-encoded IgH constant region (initially for the IgM isotype) can be exchanged through genetic mechanisms (i.e. CSR) for the exons of a different germline-encoded constant region (i.e. the IgG, IgE, or IgA isotype), located downstream of the gene, to produce antibody that maintains the original variable domain but has a different constant domain and so a different effector function profile (Figure 2) [Besmer, 2004; Cascalho, 2004; Green, 1998; Lee, 2001; Neuberger, 2003; Papavasiliou, 2002; Zarrin, 2007].

Some antibody isotypes have evolved to exploit avidity strategies that extend their antigen-binding capabilities beyond the intrinsic affinity of their antigen-binding sites indirectly by forming antibody multimers linked together by J-chains and/or disulfide bonds [Hsu, 2006]. Antibody multimerization can increase functional affinity without alteration of the Ig genes. For example, multimerization is an inherent effector function of the IgM antibody isotype, the earliest antibody expressed in an immune response, which is characteristically of low affinity [Wabl, 1999]. In serum, IgM is a pentamer with ten antigen-binding sites [Hsu, 2006; Wabl, 1999], a beneficial feature as many antigens bear repetitive epitopes.

The primary repertoire of antibody diversity is generated during B-cell development in the bone marrow, a primary lymphoid tissue, when one of each of the numerous V and J sub-exons of the IgL gene and the V, D, and J sub-exons of the IgH gene undergo recombination, directed by recombination signal sequences (RSS) that flank the sub-exons in the DNA and mediated by the recombinase products of the recombination-activating genes, RAG1 and RAG2, to yield VJ and VDJ variable region exons, respectively ([Honjo, 2004; Hsu, 2006; Neuberger, 2003] and reviewed by [Bassing, 2002]) (Figure 1). During these site-specific RAG-mediated recombination events, which contribute mainly to the formation of CDR3, other processes can randomly add or remove nucleotides from the joining ends to create unique coding joint junctions [Bassing, 2002].

In the mouse, it has been estimated that RAG-mediated recombination of the sub-exon arrays, which have been selected by evolution, coupled with junctional diversity to create the Ig gene variable region exons and the pairing of the IgH and IgL polypeptides to form the antibody variable domains has the potential to generate in the range of 10^9 to 10^{10} different binding sites for antigen recognition [Hsu, 2006; Neuberger, 2003] — a potential which may be further enhanced by conformational flexibility [Wedemayer, 1997] and presence of amino acids that promote more promiscuous binding interactions (polyreactive) [Clark, 2006]. While this vast potential for binding of antigens — far greater than that which could be germline encoded — means that there is a significant likelihood that at least some antigens of any given pathogen can be recognized by antibody [Neuberger, 2003], it also means that there is a significant likelihood that some molecules of self can be recognized by antibody, an undesirable event as it carries the potential for the development of an autoimmune response [Durandy, 2005].

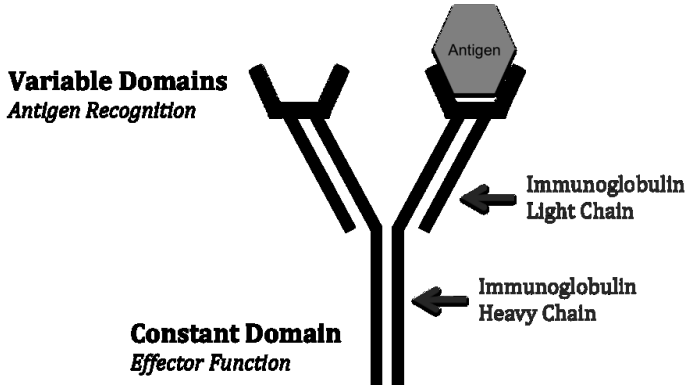
B-cell receptor-based selection of B-cells during their development leads to self-tolerance through elimination (apoptosis or anergy) or alteration (receptor editing) of B-cells bearing self-recognizing BCR [Bassing, 2002; Wabl, 1999]. To facilitate this, the process of allelic exclusion prevents each developing B-cell

from expressing more than one IgH gene and one IgL gene which ensures that each mature B-cell produces BCRs with only one type of antigen-binding site [Bassing, 2002]. Thus, the mature B-cell population is polyclonal based on BCR.

While each of the mature, naïve B-cells (i.e. those that have not encountered their cognate antigen) expresses BCRs of a single, unique type of antigen-binding site, they can express two different IgH constant domains, IgM or IgD isotypes, through alternative splicing of the IgH pre-mRNA transcript [Hsu, 2006; Ohta, 2006].

The primary repertoire of antibody diversity generated during B-cell development sets the stage for first-time encounter and recognition of a vast assortment of unspecified foreign molecules — most of which comprise a variety of epitopes [Neuberger, 2003].

A



B

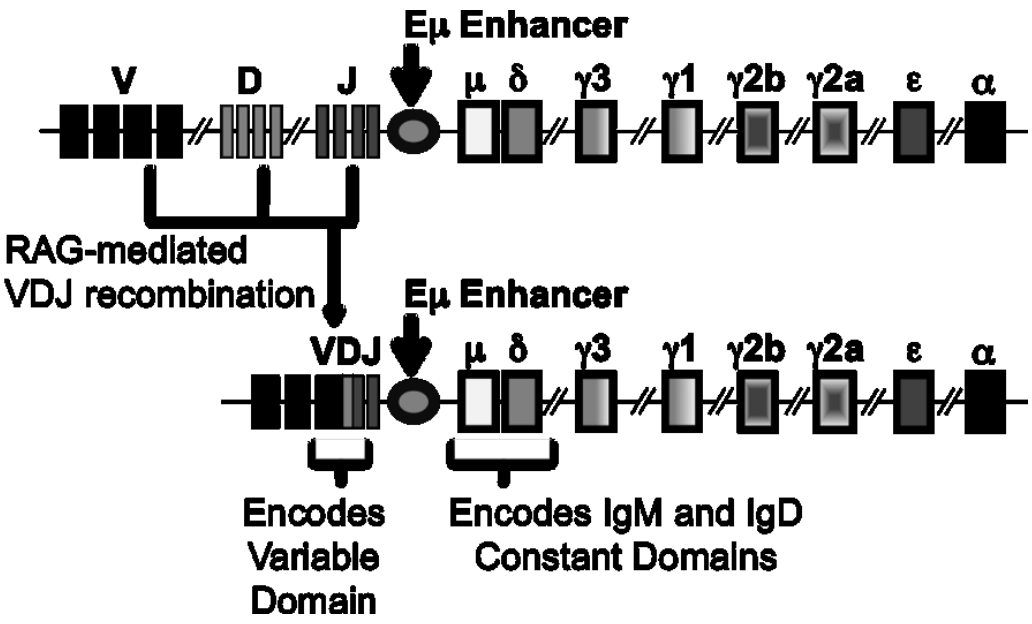


Figure 1. Mouse antibody molecule and immunoglobulin heavy chain gene. An antibody (A) is comprised of two immunoglobulin light chains and two immunoglobulin heavy chains and bears two amino-terminal variable domains and a single carboxy-terminal constant domain. The germline configuration of the immunoglobulin heavy chain gene of the mouse (B) undergoes VDJ recombination to create a complete variable region exon. Diagrams are not drawn to scale and constant region exons have been depicted as single boxes for clarity.

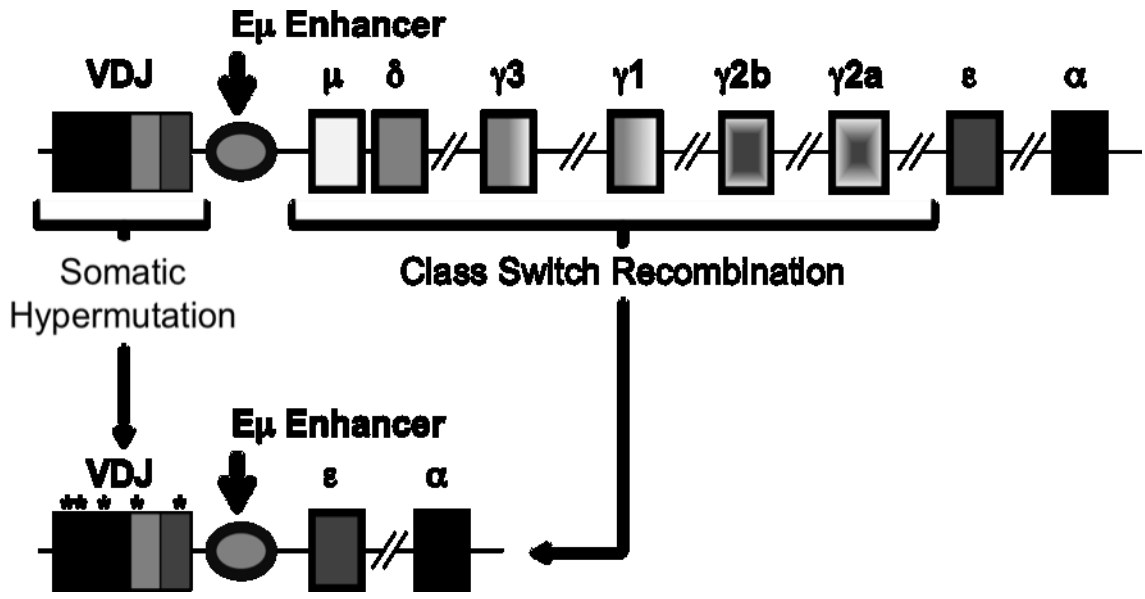


Figure 2. Somatic hypermutation and class switch recombination in the mouse immunoglobulin heavy chain gene.

During antibody affinity maturation, the IgH gene is altered by two different AID-mediated processes. Somatic hypermutation introduces non-templated point mutations (represented by asterisks) into the variable region, potentially altering the antibody variable domain. Class switch recombination exchanges the germline μ/δ constant regions for another constant region from further downstream (e.g. ε), altering the isotype and effector functions of the antibody. Note that the intronic enhancer, Eμ, is not affected by CSR due to its location. Diagrams are not drawn to scale and constant region exons have been depicted as single boxes for clarity.

1.2.2 Mammalian Secondary Antibody Repertoire Generation

Subsequent to development in the bone marrow, the mature, naïve B-cells circulate within the body as they surveil it for foreign molecules [Hsu, 2006]. An encounter with its cognate antigen leads to transduction of signals via the BCR complex that may trigger a state of cellular activation in a B-cell, especially when accompanied by coligation of other cell-surface receptors (e.g. complement receptors) and costimulatory signals from activated helper T-cells (e.g. CD154, the CD40 ligand) that are specific for the "same" antigen (i.e. a T-cell-dependent humoral immune response), which has been internalized, processed into peptides, and presented, in association with Major Histocompatibility Complex class II molecules (MHC II), by the B-cell [Batista, 2000; Benson, 2007; Durandy, 2005; Foote, 1995]. As the first step of amplification and refinement of the antibody immune response, activation initiates a new program of cellular plasticity in B-cells, distinct from that of development, during which they can undergo further Ig gene modifications before their ultimate differentiation into antibody-secreting plasma or memory B-cells [Hsu, 2006].

The strength of the signal generated through the BCR complex is linked to the BCR's affinity for the cognate antigen, which allows B-cells to discriminate between antigen interactions and respond accordingly [Batista, 2000]. In general, antigen:BCR interactions of higher affinity lead to more intense signaling through the BCR complex. But factors other than absolute antigen binding-affinity of the BCR (e.g. antigen type, epitope density) can play a role in regulating cumulative BCR signal by influencing aspects such as the clustering of receptors as a result of cross-linking or limiting diffusion of antigen during "off" stage through co-binding [Batista, 2000; Paus, 2006]. Because the BCRs of naïve B-cells tend to be of relatively low affinity and polyreactive [Clark, 2006; Foote, 1995; Green, 1998; Neuberger, 1995; Papavasiliou, 2002; Wabl, 1999; Wedemayer, 1997], the stringent requirement for other signals to accompany those of the BCR to help to synergistically meet or regulate signaling thresholds to initiate activation acts as a built-in failsafe mechanism to prevent aberrant activation in the absence of

pathogens and to reinforce that an immune response is genuinely required only when several indicators of infection are detected as being present. In addition to activation, the strength of antigen recognition and integration of co-stimulatory signals appear to also determine the fate of antigen-activated B-cells [Benson, 2007] (discussed below) which may differentiate to relatively short-lived (~ two weeks [Tarlinton, 1998]) antibody-secreting extrafollicular plasma cells directly or may relocate to primary follicles where they contribute to the formation of germinal centers [Paus, 2006], defined sites that arise in secondary lymphoid tissues (e.g. spleen and lymph nodes) after antigen challenge in T-cell-dependent humoral immune responses [Neuberger, 1995]. There, they can proliferate as centroblasts and undergo antibody affinity maturation processes that lead to the maturation (i.e. increase) in affinity for the specific antigen and/or to the substitution of the antibody constant domain [Stavnezer, 2002]. This means the B-cells that have been selectively recruited into the immune response via antigen-dependent activation may now produce antibodies that are better able to precisely recognize that specific activating antigen and to respond more appropriately to the type of pathogen that generated it. This culminates in differentiation to relatively long-lived antibody-secreting plasma cells and memory B-cells.

In the mouse/human mammalian paradigm of adaptive immunity, there are three main processes — two of which are crucially dependent on the AID enzyme — that contribute to antibody affinity maturation:

Selection – Antigen-dependent screening of B-cells that leads to positive selection of those B-cells bearing BCR of higher affinity for the antigen.

Somatic Hypermutation – Antibody affinity diversification through the introduction to the immunoglobulin gene variable region exon of non-templated point mutations initiated by the AID enzyme.

Class Switch Recombination – Effector function specialization through the recombinatorial exchange of IgH gene constant region exons through double-stranded DNA (dsDNA) breaks initiated by the AID enzyme.

Although both SHM and CSR of the Ig genes require the AID enzyme and appear to share many ubiquitously expressed trans-acting cellular factors and cis-acting genetic elements, they are independent processes with conspicuously different outcomes (Figure 2) [Brar, 2004; Durandy, 2005; Li, 2003; Maizels, 2000; Nagaoka, 2002; Nambu, 2003; Papavasiliou, 2002; Pham, 2005; Ta, 2003]. Likewise, selection, a process that results in the preferential amplification of higher-affinity antibody-expressing B-cell clones, can act independently upon the original, assorted population of antigen-specific activated B-cells and need not be preceded by either SHM or CSR [Diaz, 2002; Wabl, 1999]. However, the converse is not true as SHM and CSR, which operate at the level of the Ig genes quasi-detached from the constraints of maintaining or enhancing antibody affinity for the antigen, need to be succeeded by selection, which operates at the level of the antibody proteins (i.e. phenotypic) and is constrained by the requirement to preserve or refine their antigen-binding specificity and affinity, in order to be of benefit to the immune response [Tarlinton, 2000].

Since these processes, especially selection, appear to be greatly facilitated by the organization of the germinal center architecture, it seems that five key elements — germinal centers, selection, somatic hypermutation, class switch recombination, and AID — conspire to enhance the mammalian adaptive immune response through antibody affinity maturation.

Note that another process, gene conversion, which leads to antibody diversification through imprinting (templated recombination) of donor sequence, from a region upstream on the same chromosome (i.e. in cis) — typically pseudo-variable sub-exons — into the variable region exon, is also initiated by the AID enzyme [Arakawa, 2002; Barreto, 2005; Cascalho, 2004; Diaz, 1998; Durandy, 2006; Hsu, 2006; Li, 2003; Neuberger, 2003; Stavnezer, 2002; Yang, 2006]. Although gene conversion is known to occur in some mammals (e.g. rabbit) [Diaz, 2002; Diaz, 1998; Diaz, 2001; Papavasiliou, 2002; Ta, 2003], it has not been found to contribute to antibody diversification in either mouse or human, which are the focus of this paradigm [Diaz, 1998; Stavnezer, 2002]. As such, it will be largely ignored for the sake of brevity and clarity herein.

1.2.3 Germinal Centers Are Sites Of Antibody Affinity Maturation

Located in secondary lymphoid tissues, conventional mammalian germinal centers (Figure 3) are transient, specialized microenvironments [Diaz, 2001; Jacobs, 2001] mainly composed of relatively tight aggregations of activated B-cells, helper T-cells, and follicular dendritic cells (FDCs) [Manser, 2004]. While the totality of the germinal center response has not yet been fully elucidated, the germinal center organization appears to contribute to antibody affinity maturation by providing a sequestered site that serves both to enhance its efficiency, by bringing all the requisite participants together, and to reduce its risks, by constraining B-cells bearing antibodies that have potentially been negatively altered by AID-mediated SHM and/or CSR (i.e. preserve peripheral self-tolerance) [Allen, 2007; Manser, 2004].

Germinal centers exhibit a distinct tissue architecture that is easily seen in histological sections simply stained with Haematoxylin and Eosin (HE), although specialized detection methods can be employed to more explicitly distinguish specific cell types (e.g. differential staining for markers such as the absence of IgD staining and the binding of the lectin peanut agglutinin to germinal center B-cells [Fukita, 1998; Kelsoe, 1996; Manser, 2004; Muramatsu, 1999; Tarlinton, 1998]) and possible B-cell states within the germinal center (e.g. CXCR4 expressed on centroblasts [Allen, 2004]). Germinal centers are routinely discussed in terms of two obvious cellular zones, the dark zone and the light zone (described below) [Allen, 2004; Allen, 2007; Manser, 2004], which were originally delineated by distinct patterns in HE staining associated with differences such as cell density and type. Although a textbook germinal center organization is portrayed as roughly egg-shaped with the dark zone eccentrically located at one end — typically positioned proximate to the T-cell zone [Allen, 2004; Kelsoe, 1996] — most germinal centers don't conform strictly to this shape or polarity.

Although there are some identifiable differences in gene and protein expression among germinal center B-cells [Allen, 2004; Allen, 2007; Tarlinton,

1998], these cells are in an exceptional state of flux and consequently exhibit an increased degree of plasticity that has made it somewhat difficult to definitively pinpoint discrete cell stages among them, a task that has been further confounded by the highly dynamic nature of the germinal center environment [Allen, 2004]. Moreover, the specifics of the germinal center response itself have yet to be established [Manser, 2004] and, despite a propensity to discuss the germinal center as being a singular phenomenon, experimental findings suggest, perhaps not surprisingly, that germinal centers may vary among types of secondary lymphoid tissues (e.g. in mouse, CD23 expression by FDCs is apparently an exclusive feature of lymph node germinal centers) [Allen, 2004] and/or in response to differences in antigen (e.g. type, dose, epitope density, exposure duration, delivery mode) [Batista, 2000; Paus, 2006; Tarlinton, 1998] and/or other relevant factors (e.g. age, immune context including help and inflammation) [Allen, 2007; Manser, 2004]. That germinal center-like cellular aggregations are notorious for forming ectopically, outside of conventional secondary lymphoid tissues, as part of aberrant immune responses or other disease states, to which they contribute clones of antibody secreting B-cells that have undergone SHM and selection, certainly suggests that some measure of plasticity exists within germinal centers and/or the tissues in which they form [Manser, 2004; Park, 2005]. Nevertheless, as a point of reference for ongoing research, states have been provisionally assigned to germinal center B-cells and the following basic model (reviewed in [Kelsoe, 1996; Manser, 2004]), however contentious and possibly flawed, has long since been put forward to describe the germinal center response.

In the spleen, germinal centers arise a few days after antigen challenge as B-cells activated by antigen and cognate CD4+ T-cell help relocate to B-cell zone primary follicles where, along with accessory cells (described below), they develop germinal centers adjacent to the bordering T-cell zone [Foote, 1995; Manser, 2004; Neuberger, 1995]. The T-cell zone, which occupies the space around the blood vessels (i.e. periarteriolar lymphoid sheath) that convey antigen into the spleen, harbours T-cells that have been activated by the “same” antigen,

delivered to them by antigen presenting cells (e.g. dendritic cells), as the responding B-cells. It appears that most germinal centers are founded by only a few B-cells [Kelsoe, 1996; Neuberger, 1995] but later immigration, possibly from other germinal centers, may occur as well in some instances [Allen, 2007]. The germinal center enlarges as the B-cells proliferate [Manser, 2004], compressing the adjacent bystander cells to form a follicular mantle zone [Kelsoe, 1996] the interface of which delineates the boundary of the entire germinal center compartment [Allen, 2007]. Given that germinal center size is mainly achieved through massive proliferation from only a very few founding B-cells, it is not surprising that there is some evidence that the size of the germinal center does not depend on the absolute number of B-cells that an organism harbours [Allen, 2004].

Whether an activated B-cell differentiates into an extrafollicular plasma cell or enters into a primary follicle to participate in the germinal center response may be largely decided by the strength of its antigen recognition, with B-cells that express higher-affinity BCR generating the earlier extrafollicular plasma cell response and the residual B-cells that express lower-affinity BCR participating in the later germinal center response [Benson, 2007; Paus, 2006]. It seems that, once activation has occurred, there is no affinity-based barrier to the formation of the germinal center [Benson, 2007]. The apparently higher threshold required to become an extrafollicular B-cell suggests that a strategy has evolved in which the superior B-cells are actively recruited to contribute to immediate first-line protection while the inferior B-cells are actively recruited into an affinity improvement program — a possible fail-safe mechanism to extend peripheral tolerance in the event of aberrant activation as the extrafollicular B-cells, which may be more readily activated because of their higher-affinity BCR, do not contribute to memory. But, precisely how B-cells intrinsically determine relative differences in affinity via signaling thresholds, which are actively regulated [Tarlinton, 1998], despite considerable noise (i.e. other variable signals that support B-cell activation), is unknown [Benson, 2007; Paus, 2006].

Various stromal cells, most notably the non-phagocytic FDC, are also found in germinal centers and are believed to play important, yet generally poorly understood [Manser, 2004], roles in their formation and maintenance [Park, 2005]. At least one role of germinal center stromal cells is well established: they direct cell traffic within the germinal center via the production of chemokine gradients [Allen, 2004]. Another role may be played by the FDCs. It appears that the occurrence of FDCs, which are not bone marrow-derived, is always associated with the presence of activated B-cells [Matsumoto, 1997; Park, 2005]. In fact, these cells are not detectable in SCID mice (i.e. in the absence of B-cells and T-cells) or in the absence of certain cytokines (discussed below) but can be found after adoptive transfer of lymphocytes [Matsumoto, 1997; Park, 2005]. The FDCs are believed to make a major contribution to the germinal center response via their ability to passively trap antigen in the form of antibody-antigen clusters, known as immune complexes, which are captured and held for prolonged periods of time [Anderson, 2006; Wabl, 1999] on their surface by means of Fc receptors, which bind to the constant domains of the antibodies in the immune complexes, and complement receptors, which bind to complement (e.g. C3b) that has been deposited on the antigen, primarily as a consequence of antibody binding [Hannum, 2000; Manser, 2004]. The antigen displayed in these surface-bound immune complexes appears to be available to the germinal center B-cells and has been postulated to be important to the antibody affinity maturation process of selection (discussed below) [Manser, 2004; Wabl, 1999]. Moreover, that FDC clusters appear to arise in response to activated B-cells [Matsumoto, 1997; Park, 2005] suggests the variety of antigens that they bear may be largely restricted to those most relevant to the current response.

B-cells undergoing antibody affinity maturation in the germinal center appear to cycle between two phenotypically distinct states that characteristically separate spatially, directed by chemokine gradients, into the aforementioned dark and light zones [Allen, 2004]. The dark zone, a cell-dense region of extensive B-cell proliferation [Allen, 2004; Diaz, 1998], is composed primarily of AID-expressing centroblast B-cells [Allen, 2004; Woo, 2003], which have

downregulated their surface Ig expression [Kelsoe, 1996; Manser, 2004; Tarlinton, 2000]. The light zone is composed of surface Ig-bearing centrocyte B-cells [Tarlinton, 2000] that have entered into an apoptosis-primed state, helper CD4+ T-cells, and antigen-trapping FDCs [Allen, 2004; Manser, 2004]. Thus, the chemical alterations wrought upon the Ig genes by SHM and CSR appear to be effectively restricted to the centroblast B-cell state by their indispensable requirement for AID [Martin, 2002; Muramatsu, 2000; Revy, 2000] and the mechanical sorting process of BCR-based selection appears to be likewise effectively restricted to the centrocyte B-cell state by its requirement for available surface Ig to physically interact with antigen.

Centrocytes that have been selected (discussed below) during antibody affinity maturation in the germinal center, via survival signals generated mainly through antigen interactions (i.e. BCR-mediated acquisition and MHC II presentation) and CD4+ T-cell help (e.g. CD40/CD154 and cytokines), can progress to terminal differentiation to antibody-secreting plasma B-cells and memory B-cells that exit the germinal center, whereas those that have not been selected succumb to apoptosis [Allen, 2004; Allen, 2007; Durandy, 2005; Jacobs, 2001; Manser, 2004; Neuberger, 1995; Park, 2005; Revy, 2000; Tarlinton, 2000]. In addition to survival, the strength of these integrated signals, which are intimately linked to antibody affinity, may also dictate the centrocytes' terminal differentiation fate decisions, with those that express higher-affinity antibody differentiating to antibody-secreting plasma B-cells and those that express lower-affinity antibody differentiating to memory B-cells [Benson, 2007; Tarlinton, 2000]. This pattern of fate decisions based on stratification of BCR affinity is reminiscent of the initial fate decision of antigen-activated B-cells wherein those that express higher-affinity BCR appear to be preferentially shunted into immediate antibody production while those that express lower-affinity BCR appear to contribute to the later response [Benson, 2007; Paus, 2006]. There is also some evidence that too intense a signal through the BCR may lead to apoptosis [Kelsoe, 1996; Tarlinton, 2000], possibly one of several factors that may contribute to the endogenous affinity ceiling of antibody affinity maturation

typically being considerably lower than that which can be achieved experimentally with antibodies [Batista, 2000; Foote, 1995; Wedemayer, 1997].

It is the ability of germinal center B-cells to undergo repeated rounds of antibody affinity maturation (i.e. iterative Ig gene mutation and phenotypic selection) [Green, 1998; Tarlinton, 1998], coupled with expansive proliferation, that leads to the potential for the rapid development of a hierarchical series of antibody affinity-matured B-cells (i.e. clonal genealogies [Jacobs, 2001; Kelsoe, 1996; Rada, 2004]) [Peters, 1996] which consequently produce antibodies of progressively higher affinities [Neuberger, 1995; Papavasiliou, 2002]. Although relatively minor increases in the affinity of the antibody population can be achieved through selection alone, via preferential amplification of pre-existing higher-affinity antibody-expressing B-cell clones, SHM- and/or CSR-generated clonal heterogeneity within the Ig genes is required, in concert with selection, to achieve major increases in the affinity of the antibody population [Diaz, 2002; Papavasiliou, 2002]. Indeed, during a primary response, antibody affinity maturation in mammals can lead to antibody affinity increases of greater than 10-fold within two to three weeks [Allen, 2007], 100-fold within eight weeks, and as much as 1000-fold (rabbit) within a few months [Cain, 2002; Hsu, 1998; Jacobs, 2001], even though, under normal physiological conditions, a typical germinal center response appears to persist for only about three to several weeks [Allen, 2007; Tarlinton, 2000] — basically, the time required to clear the infection and most residual antigen and inflammatory signals.

The rapid expansion of the antigen-specific germinal center B-cell population and the concomitant apoptosis within this population due to the failure of numerous B-cells to be selected for terminal differentiation and egress to the periphery means that, in addition to the aforementioned detection methods, germinal centers can also be distinguished as sites of considerable cell proliferation and death [Manser, 2004; Tarlinton, 2000]. Moreover, the abundance of B-cell death in the germinal center results in the association of a further cell type of note, tingible body macrophages [Allen, 2004], which engulf

the debris from the apoptotic centrocytes [Allen, 2007; Durandy, 2005; Manser, 2004; Revy, 2000].

Once the infection has been quelled, the germinal centers undergo dissolution as the immune response wanes due to the lack of availability of sufficient antigen and other stimuli to sustain it; however, the precise cues that lead to germinal center dissolution are not well understood [Tarlinton, 1998].

1.2.3a Germinal Centers Do Not Wholly Determine Antibody Affinity Maturation Outcomes

Although they are transient, germinal centers are nonetheless complex and highly dynamic structures that rely on the intricate coordination of several molecular and cellular players for proper initiation, maintenance, and dissolution. A variety of naturally occurring and experimentally induced defects [Durandy, 2005; Tarlinton, 1998], a selection of which are discussed below, have been found to disturb the normal progression of germinal centers — often without alteration of the extrafollicular plasma cell pathway [Tarlinton, 1998]. Even though they differentially impact the various stages of germinal center progression [Manser, 2004], many of these defects effectively cripple the adaptive immune response as they may effectually preclude the processes of antibody affinity maturation and/or terminal differentiation to plasma and memory B-cells. In light of this and given that germinal centers are portrayed as being a core element in the mammalian paradigm of antibody affinity maturation, it is perhaps surprising that explorations of some of these disturbances have revealed not only that the presence of germinal centers does not guarantee antibody affinity maturation (e.g. AID deficiencies) but also that their disruption or absence does not inevitably wholly abrogate antibody affinity maturation (e.g. lymphotoxin-alpha deficiency) [Diaz, 2001; Tarlinton, 1998].

It seems that, despite being intimately linked, there exists a disconnect between germinal centers and antibody affinity maturation processes such that the germinal center morphology is somewhat dispensable [Green, 1998]. That is,

while it is not explicitly required for antibody affinity maturation as selection, SHM, and/or CSR have all been found to occur in its absence [Tarlinton, 1998], it appears to serve to augment antibody affinity maturation by enhancing its efficiency [Allen, 2007; Diaz, 1998] and increasing the maximum achievable antibody affinity peak. Whether in experimental systems or patients, it is the specific defect or immune context that ultimately dictates to what extent germinal centers can form and/or antibody affinity maturation can occur and explorations of these circumstances give further insight into the essential components of the germinal center response and allow for dissection of the contributions of the many different participants.

As the sole antibody-producing cells of the body, B-cells are clearly indispensable to antibody affinity maturation in the germinal center but some of their associated activities may not be as explicitly necessary to the process as they are portrayed in the germinal center model. This is borne out by well known experimental techniques used to stimulate harvested B-cells or B-cell lines *in vitro* (e.g. mitogens and signaling compounds derived from T-cells) to enter into an activated state accompanied by proliferation and possibly SHM, CSR and/or plasma differentiation in the absence of both cognate antigen:BCR complex engagement and the germinal center environment [Paus, 2006]. Thus, the generation of signals through the BCR complex is not a strict requirement to overcome thresholds for B-cell activation and some antibody affinity maturation processes, at least under certain *in vitro* conditions where other signals can be sufficiently heightened artificially (e.g. extra stimulation of receptors with signaling pathways that overlap the BCR complex signaling pathway or use of second messenger mimics that initiate later stages of the BCR complex signaling pathway to activate the same transcriptional programs) to substitute for the absence of signals through the BCR complex. But *in vivo*, this sort of indiscriminate B-cell activation would obliterate the desired specific response to antigen. Under normal physiological conditions, the cellular context of B-cells (e.g. germinal center) provides many different inhibitory and activatory signals that act cooperatively to regulate and focus their response according to the

prevailing conditions and to prevent generalized activation; indeed, it appears that harvested germinal center B-cells ($B220^{+}$, $CD19^{+}$, $GL7^{hi}$, Fas^{+} and/or IgD^{-}) don't survive for very long *ex vivo*, even as part of a mixed splenocyte population, unless they are able to suppress apoptosis (e.g. over-express Bcl-2) [Allen, 2004].

Signals generated through certain receptors on the B-cell may be able to compensate for each other (e.g. during mitogen activation) because they overlap downstream to activate many of the same transcriptional programs. Selective impairment of elements of these signaling pathways and their outcomes (e.g. transcriptional programs) reinforces their importance. For instance, germinal centers do not form and isotype switching is impaired if B-cell CD40/T-cell CD154 interactions are disrupted [Durandy, 2005; Qin, 1998] or if the inducible transcription co-activator OCA-B (a.k.a BOB.1 and OBF-1), which has been found to act synergistically downstream of the BCR complex, CD40, and IL-4 receptors [Qin, 1998], is defective in B-cells. That lack of OCA-B has a similar outcome to defects in CD40/CD154 interactions, which may have functions in addition to receptor signaling such as to maintain contact between cells in order to facilitate reciprocal interactions, suggests that signaling through CD40 is essential for B-cells to progress to their germinal center states.

Intriguingly, some defects, such as the absence of CR2, are associated with apparently enhanced antibody affinity maturation. CR2 is involved in coligation of antigen and costimulatory signaling (via recruitment of CD19 to the BCR complex [Tarlinton, 1998]) that lowers the activation threshold of naïve B-cells. Its absence results in plasma cells that express higher-affinity antibody [Tarlinton, 2000], likely because only naïve B-cells bearing BCR of relatively high affinity are able to surmount signaling thresholds of activation without it [Tarlinton, 1998]. That the mammalian adaptive immune system has evolved to utilize antigen coligation mechanisms, despite the damping effects that they appear to exert on the potential antibody affinities of the resultant plasma cells, strongly suggests that there may be considerably more benefit to recruiting a larger initial pool of

responders with a broader range of antibody affinities than there is to recruiting a smaller subset of responders with higher antibody affinities.

That germinal centers can be initiated in certain mice in the absence of T-cell help [Allen, 2007; Hannum, 2000], although they abort early due to B-cell apoptosis, and SHM can occur in germinal centers independent of T-cell help, although not efficiently [Manser, 2004], jointly suggest that T-cell help may not be entirely essential to the early stages of antibody affinity maturation in the germinal center and only becomes essential when the centrocytes require T-cell help to rescue them from apoptosis, an important component of both selection (described below) and the progression to differentiation. Indeed, blocking interactions between B-cells and T-cells in the germinal center halts the response through germinal center dissolution and disruption of differentiation [Manser, 2004].

Defects in some chemokines and their receptors have been found to result in morphological perturbations within the germinal center, suggesting that light and dark zone partitioning within the germinal center relies on B-cell responses to chemokines. Notably, upregulation of CXCR4 expression occurs in centroblasts and appears to control their movement into the dark zone, in response to CXCL12 produced by stromal cells [Allen, 2004]. This chemotaxis contributes to the establishment of dark and light zones by creating a dense cluster of centroblasts (dark zone) that displaces the other cells (light zone) within the germinal center. The distinction between the dark and light zones is abrogated by CXCR4 deficiency in B-cells but, even though they are essentially in the light zone, these CXCR4-deficient centroblasts still continue to proliferate, apparently normally [Allen, 2004]. Thus, although the dark and light zones represent segregation of B-cells in different functional states (i.e. centroblast and centrocyte) in the wild-type organism [Allen, 2004], the two zones are not themselves required for functional differences to manifest in germinal center B-cells.

Germinal center B-cells appear to be drawn into the light zone, in response to CXCL13 held on FDCs, through the constitutive expression of CXCR5 [Allen,

2004; Allen, 2007]. Unlike the CXCR4-mediated response, the CXCR5-mediated response does not seem to play a definitive role in establishing partitioning within the germinal center [Allen, 2004] but deficiency in CXCR5 or CXCL13 does lead to the generation of germinal centers of smaller size and abnormal organization [Allen, 2004]. In adoptive transfer studies, CXCR5-deficient germinal center B-cells were found to be localized to the dark zone instead of the light zone [Allen, 2004]. And CXCL13-deficiency results in the light zone being less likely to be oriented distal to the T-cell zone [Allen, 2004]. Indeed, Allen et al. [Allen, 2004] report that, in CXCL13-deficient mice, dark and light zone segregation were present in lymph node germinal centers but their arrangement was nonstandard and, in nearly one third of the germinal centers examined, the centroblasts formed a ring around a central region that housed the FDCs.

A more extreme example of germinal center disruption is seen in mice that lack the cytokine lymphotoxin-alpha (LTα). These mice do not form germinal centers, do not have detectable FDCs (i.e. no visible staining with anti-complement receptor 1 or immune complex trapping of preformed horseradish peroxidase-anti-peroxidase[rabbit] complexes), lack lymph nodes, and exhibit disruptions in the organization of the immune compartment of their spleens [Matsumoto, 1997]. Despite these seemingly significant abnormalities, they have been found to develop a high-affinity hypermutated and class-switched (IgG) antibody response to specific antigen, similar to that of wild-type mice (slightly reduced), when immunized with a “high” dose of antigen [Matsumoto, 1997; Tarlinton, 1998].

Irradiation of LTα-deficient mice followed by reconstitution with wild-type bone marrow diminished the amount of disruption in the organization of the immune compartment of their spleens and “restored” their ability to form germinal centers replete with FDC clusters [Matsumoto, 1997]. And disruption of a related but not redundant pathway via deficiency in tumor necrosis factor receptor I (TNFRI) or its ligand, tumor necrosis factor alpha (TNFα), results in abnormalities that are similar to those seen in LTα-deficiency. These mice produce peanut agglutinin-

staining cells (i.e. germinal center B-cells) after immunization but not FDCs [Tarlinton, 1998]. This suggests that the differentiation of the unknown FDC precursors to FDCs relies in part on signals initiated by their engagement of LTα and TNFα — molecules that may be supplied by activated B-cells [Matsumoto, 1997].

Interestingly, it seems that although FDCs are associated with activated B-cells [Matsumoto, 1997; Park, 2005], their numbers are not proportionate to germinal center size, which is largely determined by B-cell numbers [Revy, 2000], as AID-deficient chimeric mice, which exhibit lymphoid hyperplasia, were observed to have relatively “diffusely spread” FDCs within their germinal centers [Muramatsu, 2000]. But how FDC numbers are regulated and whether there is a defined number of FDCs per germinal center “remains unclear” [Matsumoto, 1997].

And it seems likely that the FDCs, which may not be an entirely uniform cell phenotype but instead exist as subpopulations [Park, 2005] (e.g. CD23 is found expressed on FDC only in lymph node germinal centers [Allen, 2004]), play a subtler role than just immune complex trapping [Manser, 2004]. For example, they are known to be involved in centrocyte chemotaxis [Allen, 2007], delivery of important secondary signals such as those supporting proliferation [Park, 2005], and are a more potent stimulator of B-cells *in vitro* than is soluble antigen [Park, 2005]. Indeed, a study with transgenic mice that express only membrane-bound BCR of the IgM isotype but not serum antibody found that these mice could still form germinal centers even though they should not be able to form immune complexes and antigen could not be detected on the surface of their FDCs [Hannum, 2000]. A follow-up study found that long-lived antigen-specific B-cells and memory responses were not impaired by the lack of detectable immune complexes on the FDCs [Anderson, 2006]. Taken together, the results of these studies indicate that antigen-trapping and display by FDCs do not likely play indispensable roles in the regulation of germinal center organization and progression or in B-cell maturation and support the notion that the predominant

role of FDCs in germinal centers is likely effected through other activities related to their stromal cell identity — activities which may or may not be enhanced by holding antigen for B-cells.

Thus, antigen-trapping and display may be merely adjunct functions of FDCs that serve to improve the efficiency of antigen acquisition for centrocytes, which seem to be able to otherwise acquire soluble antigen in the absence of immune complexes on FDCs [Hannum, 2000], albeit less efficiently, and enhance reciprocal interactions between the two cell types [Park, 2005]. This idea of improved efficiency of antigen acquisition may extend beyond just serving as a source of readily available antigen as studies have demonstrated that, outside of affinity, various antigen-related factors, such as density and size [Batista, 2000; Paus, 2006], can impact B-cell interactions with and responses to antigen. For example, antigen density can affect the amount of BCR engagement (e.g. promote BCR clustering and limit antigen diffusion) and thereby influence the strength of the signal [Tarlinton, 2000].

Additionally, although the antigen component is so much the focus of research into the immune complexes on FDCs that the phenomenon is commonly referred to as “antigen-trapping”, the immunoglobulin component may also have an important regulatory function in germinal centers [Hannum, 2000]. Indeed, later in the response (i.e. after class-switching has occurred), it seems that it may contribute to the transition from plasma cell production to memory cell production, possibly via the B-cell inhibitory FcγRII [Hannum, 2000; Tarlinton, 2000]. It may also participate in germinal center dissolution [Hannum, 2000], an idea supported by the fact that lymphoid hyperplasia, due to enlarged germinal centers (up to 10x) [Honjo, 2004; Maizels, 2000; Revy, 2000], appears to be an overall characteristic of hyper-IgM (HIGM) syndromes, which are characterized by a lack of immunoglobulin isotypes other than IgM, regardless of their cause [Durandy, 2006; Kavli, 2007; Minegishi, 2000], and can be diminished by passive IgG therapy [Durandy, 2005].

Intriguingly, lymphoid hyperplasia does not appear to be a feature of germinal centers in which apoptosis has been blocked in centrocytes (e.g. over-expression

of Bcl-2) as they appear “normal” in both size and morphology [Allen, 2004], although the dynamics and output of antibody affinity maturation are altered, including retention of lower-affinity antibody-expressing B-cell clones and some subtle differences in the ratio of plasma to memory B-cell populations [Tarlinton, 2000].

Overall, it appears that the germinal center scheme serves to promote antibody affinity maturation mainly through measures that enhance selection mechanisms [Diaz, 2001].

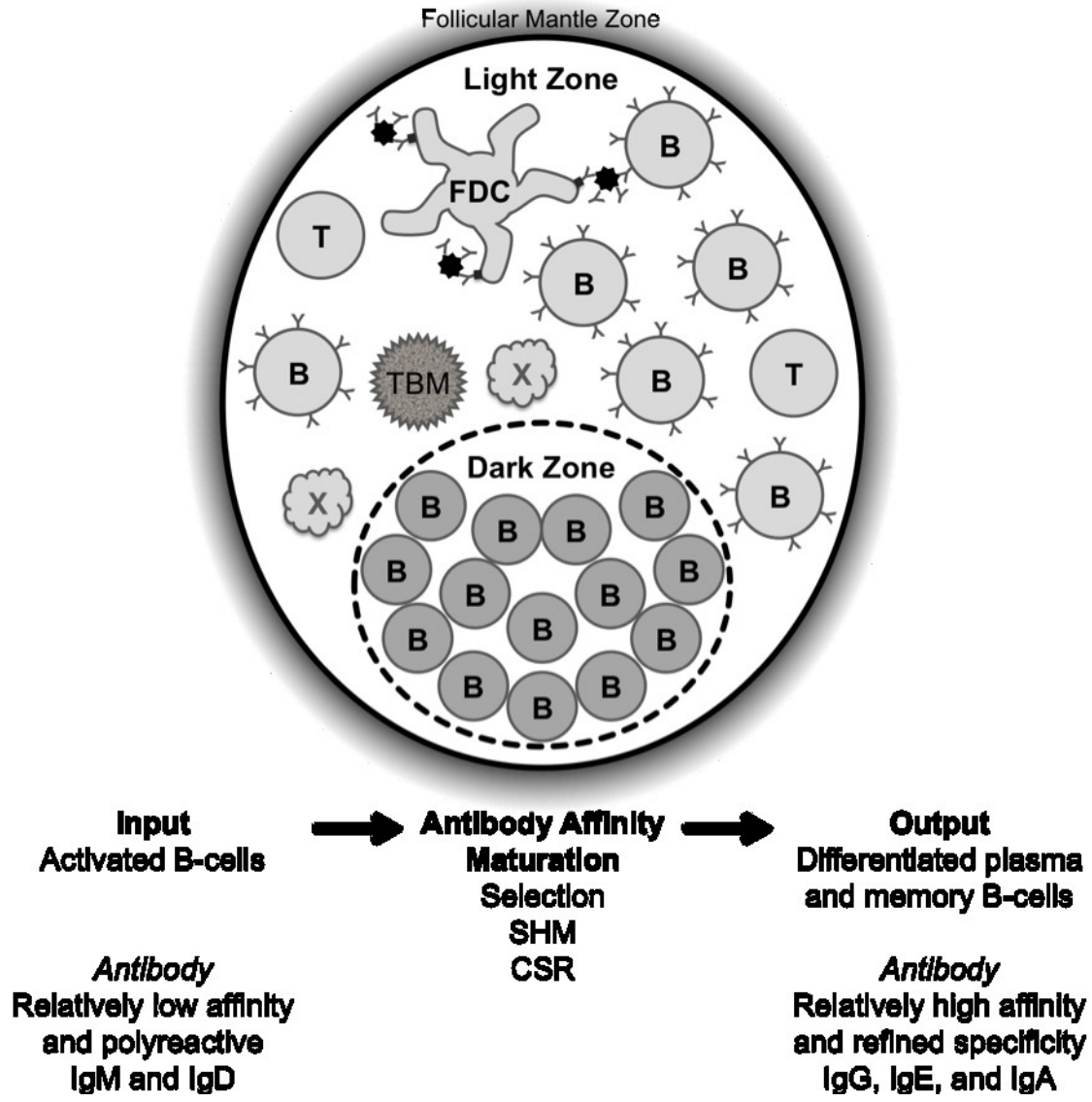


Figure 3. Simplified diagram of a conventional mammalian germinal center. Germinal centers are sites of antibody affinity maturation that develop from primary follicles seeded by a few activated B-cells. The boundary of the germinal center is delineated by the follicular mantle zone, which forms as adjacent bystander cells are compressed during germinal center expansion. The germinal center is comprised of two histologically distinct regions, the light and dark zones, that form as B-cells that cycle between two phenotypically distinct states, centrocytes and centroblasts, segregate in response to cytokine gradients. Located in the dark zone, the rapidly proliferating centroblast B-cells, which

appear to have downregulated surface Ig, express AID and undergo SHM and CSR. Located in the light zone, centrocyte B-cells, which appear to have upregulated surface Ig, undergo selection through interactions with antigen, FDCs, and helper T-cells. B-cells that fail to be selected undergo apoptosis and are consumed by tingible body macrophages. B-cell maturation culminates in the differentiation of germinal center B-cells to plasma B-cells and memory B-cells that express higher-affinity, class-switched antibodies. B = B-cells, T = T-cells, FDC = Follicular dendritic cells, X = Apoptotic B-cells, TBM = Tingible body macrophages. Adapted from [Allen, 2007; Kelsoe, 1996; Tarlinton, 2000].

1.2.4 Selection Is Crucial For Antibody Affinity Maturation

Selection features prominently in the success of adaptive immunity. This is clearly evidenced by B-cells, which are repeatedly subjected to selection events that operate predominantly upon the distinguishing characteristics of their unique BCRs. Those B-cells that have endured BCR-based selection during development in the bone marrow and then been selected, via BCR engagement of antigen, for immune activation in the periphery are further selected, yet again principally via BCR engagement of antigen, within the confines of the germinal center [Wabl, 1999].

The antigen-dependent screening of B-cells that occurs during antibody affinity maturation, within the sequestered microenvironment of the germinal center, leads to selection of antigen-specific B-cells expressing pre-existing or potentially altered (i.e. through SHM and/or CSR, discussed below), higher-affinity antibody [Diaz, 2001; Neuberger, 1995; Papavasiliou, 2002; Tarlinton, 1998]. The requirement for inclusion of a selection process in antibody affinity maturation not only ensures an increase in the overall affinity of the antibody response to a specific antigen, but also eliminates those B-cells bearing BCR that may have been adversely altered by SHM and/or CSR (e.g. a reduction in affinity for the antigen or a newly acquired or enhanced affinity for self molecules) to maintain self-tolerance [Diaz, 2001; Kelsoe, 1996; Manser, 2004; Papavasiliou, 2002].

Selection is often regarded as the final step of the three main antibody affinity maturation processes — in part because germinal center B-cells appear to be obligated to undergo selection before being allowed to differentiate and exit to the periphery. But selection is actually an ongoing process, as B-cells may repeatedly cycle between the germinal center zones [Green, 1998; Manser, 2004; Neuberger, 1995], and, because antibody affinity maturation processes are essentially discrete events that are not strictly coupled to each other [Brar, 2004] and there is no explicit obligation for SHM and/or CSR to occur, selection does not have to be preceded by either SHM or CSR [Diaz, 2002; Papavasiliou, 2002].

Additionally, because it does not rely on the expression of AID and can operate on the original assorted population of antigen-specific activated-B-cells, without the requirement for the alteration of antibody, selection is arguably the most basic of the three main antibody affinity maturation processes. However, BCR-based cellular selection subsequent to SHM and CSR refines the antibody repertoire based on the cognate antigen to yield antibodies of substantially higher affinity, whereas selection that occurs in the absence of these antibody alteration processes has significantly less potential to increase the affinity of the antibody response [Allen, 2007; Durandy, 2005; Jacobs, 2001; Neuberger, 2003].

The typical narrative given for selection in splenic germinal centers is as follows. Antigen, in the form immune complexes that have drained to the region via the blood, is captured by and held on the surface of FDCs in the light zone [Hannum, 2000; Manser, 2004]. Tethered to the surface of these cells, the trapped antigen is directly available to the apoptosis-primed centrocyte B-cells [Batista, 2000; Wabl, 1999], which may have already undergone rounds of selection, possibly accompanied by proliferation, SHM, and/or CSR [Allen, 2007]. As this antigen is available in limited quantities [Lee, 2001; Okazaki, 2003], a competitive binding situation arises among the numerous centrocyte B-cells in which only those antigen-specific B-cells bearing the highest-affinity BCR within a given germinal center manage to bind to the antigen [Honjo, 2004; Manser, 2004; Tarlinton, 1998; Wabl, 1999], an activity that also promotes other intimate interactions with the FDCs [Allen, 2007].

In this competition for limited antigen, those centrocyte B-cells that are successful at binding to the antigen internalize it via receptor-mediated endocytosis and process it into short peptides for presentation, in association with MHC II molecules, to the proximate germinal center CD4+ T-cells, which have been activated, via antigen-presenting cells, by the “same” antigen recognized by the germinal center B-cells [Foote, 1995; Manser, 2004]. This antigen-dependent interaction of the B-cells with the T-cells, itself a layer of selection through MHC restriction of the T-cells [Vallejo, 1991], leads to secondary co-stimulatory signals that result in the selection, via survival, of these

B-cells; thus, as time passes, the germinal center B-cell population becomes enriched with clones that produce antibodies of increasingly higher affinity for a specific antigen [Neuberger, 1995; Tarlinton, 1998]. In contrast, those B-cells that have not been able to acquire antigen to present to the T-cells do not receive the survival signals necessary to rescue them from their apoptosis-primed program and so succumb to apoptosis by default; consequently, B-cells bearing antibody of lower affinity for the antigen, regardless of the reason (e.g. originally low affinity or adversely altered by SHM and/or CSR [Manser, 2004]), are eliminated from the B-cell pool [Kelsoe, 1996].

B-cells that survive and exit from the antibody affinity maturation cycle can be differentiated to antibody-secreting plasma cells, which contribute to the current immune response against their cognate antigen — replacing short-lived extrafollicular antibody-secreting plasma cells — and then die off, or to relatively long-lived memory cells, which are held in reserve for a prolonged period of time against the possibility of a future re-exposure to the same antigen, an event that would re-activate these B-cells causing them to restart the antibody affinity maturation cycle described above [Allen, 2007].

Antibody affinity maturation that occurs during memory responses is operationally very similar to that seen in primary responses, with selection, SHM, and CSR occurring within germinal centers, but there are some notable differences (e.g. altered repertoire usage [Meffre, 2001]) that arise because the memory B-cells that generate these responses have differentiated from an antibody affinity-matured population of B-cells (i.e. previously hypermutated, class switched, and antigen-selected) [Anderson, 2006; Neuberger, 1995]. Due largely to the prior application of selection, memory B-cell populations typically harbour larger numbers of antigen-specific cells and express BCRs of more refined specificities and higher affinities, with respect to their cognate antigens, than did their naïve B-cell predecessor populations, attributes that have important implications for future re-exposures to antigen. First, the activation threshold is intrinsically lower in memory B-cells and their higher-affinity BCRs concomitantly generate a stronger signal upon antigen encounter [Batista, 2000], so memory B-

cells generally require less antigen and/or less intense help to become activated (i.e. relaxed stringency). This enhanced sensitivity, coupled with the higher initial number of antigen-specific cells, most of which are already switched to the optimal effector isotype, contributes to a more rapid response against proven threats that is typically of shorter duration than primary responses. Second, the significantly higher initial antibody affinity of memory responses allows antibody affinity maturation in secondary and subsequent memory responses to generate an even higher affinity peak than was achievable in the primary or previous memory responses, even though memory responses are, as noted, typically of shorter duration than the primary response.

This narrative of selection presents a simplified overview of events, but, as with the basic model for germinal centers presented above, selection is actually quite complex and there is ongoing discussion about the precise details and selective pressures involved. A main point of dissension from the given narrative focuses on the presumed importance of germinal centers. Although selection is an integral component of the germinal center response, it seems that the germinal center environment itself is not essential to selection as some antibody affinity maturation does occur in the absence of germinal centers in mice [Allen, 2007; Green, 1998; Matsumoto, 1997].

Since some antibody affinity maturation can occur in the absence of germinal centers, the roles of the individual elements within the germinal center may be likewise inflated in importance in this narrative of selection. For instance, the importance of antigen-trapping by FDCs to selection in the germinal center may be overstated because, even though the acquisition of antigen is essential for germinal center B-cell survival and selection, the absence of immune complexes (i.e. antigen) on the surface of FDCs does not seem to impair selection, suggesting that the germinal center B-cells are sufficiently able to acquire antigen in other ways (e.g. soluble antigen) [Hannum, 2000].

Limited availability of antigen, irrespective of how it is acquired, is generally recognized as a significant antibody affinity-enhancing factor in germinal center selection via competitive binding among B-cells. Nevertheless, that the

“stringency” of antibody affinity maturation is not diminished when FDCs are laden with an abundance of immune complexes (i.e. excessive antigen) [Manser, 2004] suggests that other essential components of selection may also be in relatively short supply in the germinal center. The selection model typically asserts that those germinal center B-cells that are able to outcompete their peers, and so acquire antigen, will then go on to receive survival signals from the germinal center T-cells. But the acquisition of antigen might not automatically guarantee B-cells the requisite access to T-cells. Allen et al. [Allen, 2007] suggest that, because B-cell/T-cell interactions are essentially one-to-one due to the polarization of receptors to the interface, the availability of T-cell help may also be a critical limiting factor in germinal center selection, adding another layer of competition among those B-cells that have managed to acquire antigen.

Similarly, the role of immune complexes in selection may be more nuanced than merely providing a source of antigen. As antibody affinity maturation progresses and higher-affinity antibodies are assimilated into the immune complexes, binding to epitopes increasingly strongly, these higher-affinity antibodies may become competitive with the B-cells [Tarlinton, 2000], potentially forcing the response to shift focus to other epitopes.

While there is ongoing debate about the extent to which centrocytes are actually pre-primed for apoptosis, there is no refuting the fact that they do undergo apoptosis in large numbers within the germinal center. The important role that apoptosis plays in selection is highlighted by the fact that there is a significant reduction in average antibody affinity of plasma cells and serum antibody in mice that over-express a suppressor of apoptosis (bcl-xL) [Tarlinton, 2000]. Thus, selection within the confines of the germinal center provides for rapid increases in antibody affinity by an elegant system that provides checks and balances that help to accomplish this not only by the preferential amplification of higher-affinity antibody-expressing B-cell clones but also by the eradication of B-cells that express relatively lower-affinity antibody which would dilute the serum antibody response and so diminish the potential ultimate affinity maxima of the serum antibody response.

It appears that functionally, germinal center selection approximates evolution [Clark, 2006; Wabl, 1999] in that application of selective pressure — BCR must adequately bind to antigen to receive help — to a population leads to selection via differential survival due to differences in a heritable trait — Ig genes that have been diversified during B-cell development and, possibly, through AID-mediated modification (i.e. SHM and CSR) in the germinal center B-cells.

1.2.5 Somatic Hypermutation Can Diversify Antibody Affinity

Somatic hypermutation, which occurs during the centroblast state of B-cell maturation in the germinal center [Martin, 2002; Neuberger, 1995; Poltoratsky, 2001; Woo, 2003], is the AID-mediated alteration of the expressed rearranged Ig genes by introduction of non-templated point mutations into the variable region exons, which may result in the diversification of antibody affinity and/or specificity in a way that is beneficial, neutral, or detrimental to the immediate immune response and/or the organism [Cascalho, 2004; Chaudhuri, 2004; Diaz, 2002; Green, 1998; Jacobs, 2001; Neuberger, 2003; Papavasiliou, 2002; Yang, 2006]. This SHM-generated diversity within the Ig genes of the germinal center B-cell population (i.e. secondary antibody repertoire) significantly enhances antibody affinity maturation by providing potentially improved antigen-specific antibody for selection to act upon [Diaz, 2002; Durandy, 2006; Jacobs, 2001; Neuberger, 2003; Wabl, 1999].

Although the normal background spontaneous mutation rate of the genome in non-AID-expressing cells is approximately 10^{-9} mutations/base pair/generation, the SHM rate in Ig genes in AID-expressing B-cells is about 10^{-3} mutations/base pair/generation — one million-fold higher than the background spontaneous mutation rate [Bachl, 2001; Bransteitter, 2006; Diaz, 2002; Durandy, 2006; Durandy, 2005; Green, 1998; Jacobs, 2001; Kinoshita, 2006; Neuberger, 1995; Pham, 2005; Poltoratsky, 2000; Rogozin, 2004; Wabl, 1999; Wang, 2004]. It is this high rate of mutation, which allows for significant changes to occur in the variable regions of the Ig genes in a relatively short period of time, that, when

coupled with selection, helps to drive the rapid microevolution of the antibody population during the immune response to its cognate antigen [Durandy, 2006; Wabl, 1999]. Indeed, studies of humans and animals with naturally or artificially diminished or absent SHM have revealed that SHM is an invaluable mechanism for antigen-driven antibody repertoire diversification [Durandy, 2005; Meffre, 2001] without which antibody affinity maturation can achieve only nominal increases [Diaz, 2002], imperiling the defective individuals' survival [Durandy, 2006].

Somatic hypermutation occurs predominantly as single nucleotide substitutions [Diaz, 1998; Poltoratsky, 2000], although rare small insertions (~1% of mutations) and deletions have been found to occur (~4-10% of mutations) [Clark, 2006; Diaz, 1998; Durandy, 2006; Durandy, 2005; Gordon, 2003; Jacobs, 2001; Neuberger, 1995; Papavasiliou, 2002; Peters, 1996; Rada, 2004; Rada, 2002; Stavnezer, 2002; Yoshikawa, 2002]. Yet the Ig genes of antibody affinity-matured B-cells are found to have accrued multiple mutations, typically distributed throughout the VDJ sequence [Green, 1998]. The average mutation frequency has been found to increase as the response progresses *in vivo* [Green, 1998; Tarlinton, 2000] and also through successive activations *in vitro* [Tarlinton, 1998]. This is the cumulative result of rounds of iterative mutation and functional selection [Clark, 2006; Diaz, 2002], accompanied by proliferation, which lead to the development of B-cell lineages (i.e. clonal heterogeneity), within a germinal center, that may express increasingly higher-affinity antibody with each generation and so exceed the affinity limits of the original responding B-cell population [Neuberger, 1995].

The occurrence of somatic mutation within the germline of an individual multicellular organism is generally very undesirable to that individual as it can lead to genome instability, which can increase the risk of cellular dysregulation that can lead to morbidity and mortality (e.g. cancers) [Besmer, 2004; Green, 1998; Okazaki, 2003; Pham, 2005; Wabl, 1999]. The importance of limiting genome instability is evident not only from the number of mechanisms that have

evolved to prevent and/or repair genetic mutations and the considerable resources dedicated to this end, but also from the problems that occur when these mechanisms fail [Barreto, 2005; Green, 1998; Jacobs, 2001; Pham, 2005; Wabl, 1999]. Despite the potentially dire consequences, the evolution and maintenance of mechanisms that facilitate Ig gene SHM suggest that somatic mutation in this context is also very desirable, if only under certain tightly controlled conditions.

Conventional SHM has evolved within these constraints and meets the obligation for strict regulation by employing a strategy of multiple layers of regulation to allow for precise control [Green, 1998]. It is a complex process that requires both cis-acting genetic elements (e.g. promoter and enhancer sequences) and trans-acting cellular factors (e.g. AID and DNA damage tolerance and repair enzymes) to occur [Diaz, 2002; Wabl, 1999] and is restricted to defined portions of specific genes during a specialized state of a distinct cell type — the variable regions of the expressed Ig genes of centroblast B-cells [Martin, 2002; Pham, 2005; Woo, 2003] which populate the dark zone of the germinal center. Despite strict regulation, some non-Ig genes (e.g. Bcl-6 gene [Aoufouchi, 2008; Bachl, 2001; Brar, 2004; Diaz, 2002; Durandy, 2005; Jacobs, 2001; Michael, 2002; Papavasiliou, 2002; Poltoratsky, 2000; Wabl, 1999; Wang, 2004; Zhang, 2001] and BCR associated proteins CD79α and β [Gordon, 2003]) do appear to be frequent targets of a small amount of SHM (1/100 to 1/10 of Ig SHM frequency [Bransteitter, 2006; Jacobs, 2001]) in both normal activated B-cells [Bransteitter, 2006; Chaudhuri, 2004; Larijani, 2005; Michael, 2003] and, more commonly, in cancerous B-cells where it is not always clear whether aberrant SHM is the cause or a result of the cancerous (i.e. dysregulated) state [Barreto, 2005; Besmer, 2004; Brar, 2004; Cascalho, 2004; Gordon, 2003; Green, 1998; Harris, 2002; Kinoshita, 2006; Neuberger, 2003; Okazaki, 2003; Papavasiliou, 2002; Pham, 2005; Wabl, 1999; Wang, 2004]. The increased probability of mutation in particular non-Ig genes may reflect a lack of finite control of SHM, possibly due to overlap with Ig genes in the use of cis-acting genetic elements (i.e. regulatory sequences) that direct SHM to them [Diaz,

2002; Gordon, 2003; Michael, 2002; Michael, 2003; Wang, 2004] or to other yet to be identified reasons, but the risks of adverse outcomes from apparently aberrant SHM targeting may be mitigated by rigorous selection [Gordon, 2003; Wang, 2004].

The cell type and state restrictions of SHM [Green, 1998; Papavasiliou, 2002; Wabl, 1999] are now known to be primarily dictated by the expression of the recently discovered AID mutator enzyme, the B-cell-specific trans-acting cellular factor that is indispensable for SHM as it is responsible for initiating the mutational process by deaminating C to U in the DNA of the variable region exons of the Ig genes (discussed below) [Honjo, 2004; Muramatsu, 1999; Muto, 2000; Pham, 2005]. The means by which SHM is restricted to the variable region exons of the expressed IgL and IgH genes are less clearly understood [Green, 1998], but appear to rely on choreographed access to the target genes and recruitment of AID and DNA damage tolerance and repair enzymes [Papavasiliou, 2002; Woo, 2003].

Access to the IgH gene is a requirement for SHM and represents a key layer of control [Poltoratsky, 2000]. Woo et al. [Woo, 2003] identified signatures of chromatin structure reorganization (i.e. hyperacetylation of histones) in the variable region exon, but not in the constant region exons, of the IgH gene in a B-cell line that had been stimulated to induce SHM; further to this, they were able to “extend” the occurrence of mutations into the constant region exons by chemically (i.e. histone deacetylase inhibitor) forcing the reorganization of the chromatin in the constant region during SHM. Similarly, Bachl et al. [Bachl, 2001] showed that the presence of a histone deacetylase inhibitor resulted in an increase in mutation rate in a GFP transgene.

Chromatin remodeling allows for transcription and it seems that, ultimately, it is transcription that can provide both access to the gene (e.g. transcription bubble [Ramiro, 2003]) and link it to the mutational machinery (e.g. AID and DNA damage tolerance and repair enzymes) [Bachl, 2001; Barreto, 2005; Diaz, 2002; Imai, 2003]. Indeed, SHM does not occur in the absence of transcription [Yang, 2006] and there is a positive correlation between the level of sterile transcription

(i.e. increased access) in the IgH gene and the frequency of SHM therein [Bachl, 2001; Besmer, 2004; Bransteitter, 2006; Fukita, 1998; Green, 1998; Jacobs, 2001; Peters, 1996; Ramiro, 2003; Woo, 2003; Yoshikawa, 2002]. Furthermore, the insertion of a transcriptionally active promoter into the J-C intron 5' of the constant region exons in IgH and IgL transgenes is sufficient to lead to the occurrence of mutation in the constant region [Besmer, 2004; Green, 1998; Jacobs, 2001; Kelsoe, 1996; Michael, 2002; Michael, 2003; Peters, 1996; Yang, 2006]. That an absence of mutation is observed over several hundred bases (~ 300-500 bp) before the inserted promoter supports that the mutation observed in the constant region is due to the effects of the inserted promoter [Peters, 1996].

In the IgH gene, the area of occurrence of SHM is ultimately bounded, quite abruptly but not absolutely, on the 5' end by the transcriptional start site [Kelsoe, 1996; Peters, 1996; Rada, 2001] and, more diffusely, on the 3' end by the J-C intron, which houses the intronic enhancers and matrix attachment region (MAR), such that it does not occur in the exons of the constant regions [Diaz, 2002; Neuberger, 1995; Poltoratsky, 2000; Wabl, 1999; Wang, 2004]. Moreover, the distribution of SHM within this region is skewed with the greatest amount of mutation occurring toward the 5' end of the region — beginning approximately 200 bp 3' of the promoter (i.e. within the leader intron) [Diaz, 2002; Michael, 2002], a feature which is actually sequence independent [Rada, 2001] — and steeply declining at approximately 500 bp downstream of the promoter [Bransteitter, 2004] before further diminishing on approach to the 3' SHM boundary [Rada, 2001] approximately 1.5 to 2 kb downstream of the transcriptional start site [Bachl, 2001; Bransteitter, 2006; Chaudhuri, 2004; Diaz, 2002; Gordon, 2003; Green, 1998; Michael, 2002; Peters, 1996; Pham, 2005; Rada, 2001; Wabl, 1999]. This skewing means that not only is SHM focused on the VDJ exon [Green, 1998; Ramiro, 2003] but also that the more 5' CDR1 and CDR2, which aren't impacted by junctional diversity during VDJ recombination, are subject to a disproportionate amount of SHM relative to the more 3' CDR3 [Neuberger, 1995; Rada, 2001]. Thus, it appears that SHM is somehow restricted to the variable region of the IgH gene in such way that it does not alter either the

locus control elements or the constant region exons, both of which are crucial to the continued expression of functional IgH [Neuberger, 1995].

While the requirement for sterile transcription to provide access to the Ig gene quite handily explains why the abrupt 5' boundary of SHM is the transcription start site, the precise reasons for the more diffuse 3' boundary located 1.5 to 2 kb downstream from the start site and proximate to or within the intron are somewhat more enigmatic [Diaz, 2002; Green, 1998; Peters, 1996; Rada, 2001]. The lack of chromatin reorganization further downstream, partly due to the presence of the MAR in the intron, may contribute to the 3' boundary, but the intronic enhancers may also serve a fundamental role in delineating these boundaries and shielding the constant region exons from mutation. This might be effected as a consequence, incidental or otherwise, of their presumed primary role in sterile transcription (i.e. access) and/or their postulated secondary role(s) in the targeting of SHM to the variable region exon (e.g. recruitment of AID and DNA damage tolerance and repair enzymes) [Diaz, 2002] such that transcription factors physically bound to the enhancers act as a physical block that prevents transcription, and so SHM, from proceeding further downstream [Peters, 1996]. Alternatively, distance alone may establish the 3' boundary [Green, 1998; Rada, 2001] and prevent SHM from proceeding into the constant region exons [Neuberger, 1995; Peters, 1996] since SHM that aberrantly occurs in non-Ig genes appears to be likewise confined to within 2 kb of the promoter [Pham, 2005].

Heterologous sequences (i.e. non-Ig) can act as SHM targets [Bachl, 2001; Chaudhuri, 2004; Diaz, 2002; Green, 1998; Kelsoe, 1996; Michael, 2002; Muramatsu, 2000; Peters, 1996; Poltoratsky, 2000; Ramiro, 2003; Wabl, 1999; Wang, 2004; Wang, 2004]. There are some differences in SHM between transgenes and the endogenous Ig loci (e.g. rate of mutation) but, even when transgenes contain heterologous sequence, these variations likely result from differences in promoter or enhancer elements or effects related to the position of integration [Aoufouchi, 2008; Bachl, 2001; Bransteitter, 2006; Rada, 2001; Yang, 2006]. Hypermutation studies of promoters and enhancers have found that these

sequences can profoundly influence SHM [Wang, 2004], but have not yet clearly identified an absolute requirement for specific cis-acting genetic elements within the Ig loci [Bachl, 2001; Bransteitter, 2006; Diaz, 2002; Diaz, 1998; Fukita, 1998; Green, 1998; Jacobs, 2001; Michael, 2002; Michael, 2003; Neuberger, 1995; Papavasiliou, 2002; Peters, 1996; Ramiro, 2003; Wabl, 1999; Woo, 2003; Yang, 2006]. That replacement of the IgH gene promoter with a heterologous promoter often has little effect on SHM suggests that its primary role in SHM may be to provide access to the gene through the initiation of sterile transcription [Green, 1998; Jacobs, 2001; Michael, 2002; Neuberger, 1995; Wabl, 1999]. However, that it may also make other contributions is highlighted by a promoter exchange study in DT40 IgL that identified a promoter that resulted in a reduction in mutational events (gene conversion and SHM) despite driving increased transcription [Yang, 2006]. The precise role of the IgH intronic enhancer/MAR sequence in IgH gene SHM appears to be more enigmatic than that of the promoter [Bachl, 2001; Bransteitter, 2006]. It appears to be partially dispensible, although SHM of the variable region exon is increased by its presence, especially in the correct orientation (n.b. orientation does not affect transcription) [Kelsoe, 1996; Papavasiliou, 2002; Peters, 1996; Wabl, 1999], and SHM of a GFP reporter distributed throughout the genome is actually reduced by its presence in one experiment [Wang, 2004] and increased in another [Bachl, 2001]. And the scenario is similar in IgL [Yang, 2006]. It has been suggested that the position of the enhancer (i.e. distance from promoter) may be an important factor for its role in SHM [Besmer, 2004], but this assertion is contradicted by data from various studies [Jacobs, 2001].

That SHM is also increased by the presence of sequence for transcription factor binding that doesn't alter transcription [Michael, 2003] and the removal of some transcriptional enhancer elements eliminates SHM but does not appear to significantly diminish transcription levels [Kelsoe, 1996] supports the idea that these sequences may play multiple roles in SHM [Besmer, 2004; Bransteitter, 2006; Jacobs, 2001; Neuberger, 1995; Yang, 2006]. For example, in addition to providing access through transcription, they may anchor recruitment protein(s)

for SHM machinery [Bransteitter, 2006; Jacobs, 2001; Michael, 2003; Yang, 2006].

An attempt by Fukita et al. [Fukita, 1998; Jacobs, 2001] to resolve the contributions of recruitment versus transcription to SHM *in vivo* by comparing mutation in an IgH passenger transgene without selective pressure, due to the presence of a stop codon in the leader sequence, driven by either pol II or pol I promoter in mice did not show any difference in mutation frequency when normalized to transcription levels. But they did have some reservations, based on the pol I transcripts being processed like pol II transcripts, that pol II may have been able to transcribe the pol I transgenes [Fukita, 1998; Papavasiliou, 2002], possibly via cryptic promoter sequences as these have been found to be active in some Ig sequences [Papavasiliou, 2002; Yang, 2006].

In addition to the cis-acting genetic elements that appear to contribute to SHM regulation by delimiting boundaries and controlling access and recruitment of SHM machinery to the Ig genes, the primary coding sequence itself also plays a fundamental role in directing SHM outcomes within the Ig genes [Diaz, 2002]. Although SHM within the Ig genes is usually described as being a stochastic process, this is not an entirely accurate description as it occurs preferentially at consensus sequence motifs, so-called “hotspots”, which are generally recognized as R<u>G</u>YW, and its inverse complement WR<u>C</u>Y, and W<u>A,</u> and its inverse complement <u>T</u>W (underlined base is the mutationally “hot” site in each motif), but core motifs (e.g. WRC) and other restricted subsets (e.g. AGC/T) are also commonly cited [Clark, 2006; Green, 1998; Martin, 2002; Pham, 2003; Poltoratsky, 2000; Ramiro, 2004; Zhang, 2001][Bransteitter, 2006; Cascalho, 2004; Chaudhuri, 2004; Diaz, 2002; Diaz, 2001; Jacobs, 2001; Michael, 2002; Milstein, 1998; Neuberger, 1995; Papavasiliou, 2002; Pham, 2005; Poltoratsky, 2001; Poltoratsky, 2004; Rogozin, 2004; Wabl, 1999; Yoshikawa, 2002]. The bases other than that identified as the mutationally “hot” site are also subject to mutation, but to a lesser degree (e.g. mutability within RGYW is G>R>Y>W) [Jacobs, 2001]. And, while sterile transcription of the Ig gene appears to be essential to provide access for SHM and its levels positively correlate with

mutation frequency [Bachl, 2001], the mutability of a particular sequence that serves as an SHM target is ultimately determined by its hotspot motif content [Bransteitter, 2003; Chaudhuri, 2004; Michael, 2002; Rada, 2001]. Indeed, hotspot motifs account for about 50% of mutations with the rest occurring in coldspots (SYC), sequence motifs of lower mutability, and neutral places [Michael, 2002], the mutability of which can be influenced by proximity to hotspots [Bransteitter, 2004; Chaudhuri, 2004; Clark, 2006; Green, 1998; Milstein, 1998; Neuberger, 1995; Zhang, 2001]. Furthermore, there is a hierarchy of sequence preference within these sets of motifs [Besmer, 2004; Green, 1998; Papavasiliou, 2002; Pham, 2005; Rada, 2001; Zarrin, 2004], owing largely to the presence of more or less favourable di and/or trinucleotide combinations within the core motif [Bransteitter, 2004; Green, 1998; Michael, 2002; Milstein, 1998], and among them as not all of the motifs within a set are targeted equally and alteration of a hotspot in the target sequence can generate changes in the favourability of hotspots elsewhere in the sequence [Green, 1998; Jacobs, 2001; Michael, 2002; Papavasiliou, 2002]. While the distance from the promoter has a significant impact on the mutation frequency, it does not appear to alter the SHM hotspot “pattern” [Rada, 2001]. Thus, the sequence can affect the frequency of SHM and influence its distribution [Besmer, 2004; McBride, 2004; Michael, 2002].

This differential mutability of DNA sequence is largely the consequence of two main factors: 1) the DNA target preferences of the mutator enzyme AID, which favours deaminating C to U in consensus RGYW/WRCY motifs [Martin, 2002]; and 2) the idiosyncrasies of replicative DNA polymerases and DNA damage tolerance and repair enzymes employed to resolve the pre-mutagenic U DNA lesion [Martin, 2002], which also affect the spectrum of mutations [Sohail, 2003; Wabl, 1999; Wang, 2004] and generate secondary mutation in A:T [Di Noia, 2002; Michael, 2002; Poltoratsky, 2000] (i.e. two phases of mutation contribute to observed mutational outcomes in SHM [Neuberger, 2005; Rada, 2002]; discussed below). Sequence context can exert a substantial impact on SHM through these factors [Milstein, 1998; Neuberger, 2005; Neuberger, 1995;

Poltoratsky, 2001], but it also influences mutability and contributes to antibody affinity maturation in other, complementary ways.

One of these ways relies on the partial disconnect between DNA mutation and amino acid change, a consequence of codon degeneracy, wherein mutation at the DNA level may result in a change at the amino acid level (i.e. replacement or non-synonymous mutation) or not (i.e. silent or synonymous mutation) [Clark, 2006]. This feature permits considerable DNA sequence flexibility while still maintaining protein sequence integrity (e.g. stability) [Clark, 2006]. Moreover, the position of the DNA mutation within the codon has a differential impact on amino acid coding such that mutations in the third base of the codon (i.e. wobble position) are more likely to result in silent mutations than DNA mutations in the first or second base position. Because mutations in certain DNA codons are more likely to result in amino acid changes in the protein and, due to targeting and repair specificities of the SHM machinery, certain codons are more likely to be mutated during SHM, preferential utilization of codons (i.e. codon bias) can be used to enhance or restrict mutation in the sequences [Conticello, 2005; Jacobs, 2001; Neuberger, 1995]. Indeed, sequence analyses indicate that Ig gene evolution appears to have exploited codon degeneracy, coupled with repair proclivities and the flexibility of the SHM target consensus sequences, to exert some control on the regional effectiveness of SHM, favouring not only the generation of SHM hotspots in the CDRs and coldspots in the FRs but also an increased probability that mutations in the CDRs are replacement and in the FRs are silent [Clark, 2006; Conticello, 2005; Diaz, 1998; Green, 1998; Milstein, 1998; Neuberger, 1995; Rada, 2001].

Although it does not account for the importance of position within the codon, an examination of the outcomes of all possible DNA mutations within the codons reveals that transitions are more likely than transversions to result in silent mutations [Diaz, 2001; Neuberger, 1995]. This is a beneficial bias that, in the context of DNA damage tolerance and repair, can help to maintain the integrity of the genome under normal circumstances [Diaz, 1998]. Biased nucleotide exchange has been found to occur in SHM [Kelsoe, 1996; Martin, 2002] with

transition mutations occurring more often than transversion mutations (approximately 3:1) when no selection is applied [Clark, 2006; Diaz, 2002; Diaz, 1998; Diaz, 2001; Green, 1998; Jacobs, 2001; Kelsoe, 1996; Martin, 2002; Neuberger, 1995; Pham, 2005; Poltoratsky, 2001; Yoshikawa, 2002] despite transversions technically being a more likely occurrence (2:1) if mutational events are entirely random (i.e. no preferences exerted by mutational mechanism) [Di Noia, 2002; Rada, 2002]. That transitions are more likely than transversions to occur during SHM is not surprising because replication or resolution of the pre-mutagenic U DNA lesions by the replicative DNA polymerases and ubiquitous DNA damage tolerance and repair enzymes, which generate the mutational spectrum in SHM (discussed below), is not random [Di Noia, 2002; Neuberger, 2005]. These enzymes, which maintain genome stability, exhibit a propensity to reduce mutation in both the genome and the proteome. The co-option of systems that are predisposed to reduce mutation frequency and favour transitions as a means to generate SHM appears to be a rather inefficient way to effect change.

Inefficiencies in SHM can be compensated for by the application of phenotypic selection. The presence of functional selection can mask the primary mutational process [Diaz, 2002; Poltoratsky, 2000] and skew the metrics by which SHM is assessed (e.g. mutational frequency, distribution, spectrum, and rate), but it also gives rise to other signatures of antibody affinity maturation. For example, a hallmark of antibody affinity-matured B-cells (i.e. mutated and selected) is a higher ratio of amino acid replacement to silent mutations (R/S) in the CDRs and a lower R/S in the FRs beyond what would be expected from random mutation despite the aforementioned sequence biases [Allen, 2007; Diaz, 1998], whereas the pattern of SHM in unselected cells exhibits a more uniform spread of mutations [Kelsoe, 1996].

Diversification of antibody affinity by SHM is a complex process involving AID, replicative DNA polymerases, DNA damage tolerance and repair enzymes (discussed below), and strictly regulated access to optimized gene sequences. Despite the inherent risks of the mutational process, diversification of antibody by SHM is essential to the production of rapid increases in antibody affinity as it has

the potential to generate significantly enhanced substrate for functional selection in the germinal center. It is complemented by class switch recombination, a mechanistically related but genetically distinct process that enhances response specificity [Cascalho, 2004; Pham, 2005].

1.2.6 Class Switch Recombination Alters Antibody Effector Function

Class switch recombination, which occurs during the centroblast state of B-cell maturation in the germinal center, is the AID-mediated exchange, through region-specific deletional genetic recombination, of the immunoglobulin heavy chain gene constant region from μ and δ to γ, ε, or α which encode the IgM, IgD, IgG, IgE, and IgA isotypes, respectively [Bransteitter, 2006; Chaudhuri, 2004; Durandy, 2005; Honjo, 2004; Lee, 2001; Maizels, 2000; Muramatsu, 2000; Nambu, 2003; Neuberger, 2003; Pham, 2005; Rada, 2002; Ramiro, 2003; Tashiro, 2001; Yu, 2005]. This change in constant region modifies the effector function of the expressed antibody to refine (i.e. mature) the immune response to the type of pathogen that produced the antigen to which the B-cell is responding while maintaining specificity of the variable region [Bransteitter, 2006; Cascalho, 2004; Durandy, 2006; Durandy, 2005; Ito, 2004; Lee, 2001; Li, 2003; Maizels, 2000; Papavasiliou, 2002; Ramiro, 2003; Stavnezer, 2002; Tashiro, 2001].

Remarkably, this recombination event is distinct from the RAG-mediated recombinations that occur in the Ig loci during B-cell development and instead shares several features with SHM [Li, 2003; Nambu, 2003]. Both CSR and SHM are restricted to the same specialized state of the same distinct cell type and require both cis-acting genetic elements and trans-acting cellular factors to provide access and to target AID-mediated events to defined portions of a specific gene [Barreto, 2005; Nagaoka, 2002]. But unlike SHM, which occurs in the variable region of both expressed IgL and IgH genes, CSR occurs only in the constant region of the expressed IgH gene [Nagaoka, 2002].

Although they share some elements (e.g. AID, UNG, and transcription), CSR and SHM have distinct outcomes [Green, 1998; Kenter, 2004; Li, 2003] and the

differences between them appear to allow them to be regulated independently [Durandy, 2006; Durandy, 2005; Nagaoka, 2002; Nambu, 2003; Pham, 2005]. And, despite strict regulation, as in SHM, some non-Ig sequences (e.g. c-Myc [Aoufouchi, 2008; Kinoshita, 2006; Zarrin, 2004]) do appear to be sporadic targets of recombination during CSR, which results in translocation of potentially oncogenic genes to a position downstream of the IgH gene promoter and intronic enhancer [Gordon, 2003; Jacobs, 2001; Kinoshita, 2006; Neuberger, 2003; Papavasiliou, 2002; Pham, 2005]. Intriguingly, it appears that interchromosomal isotype switch does occur occasionally [Hannum, 2000]. However, that the benefits of CSR outweigh the risks can be seen from patients with HIGM spectrum disorders, which can vary in severity and may differentially affect isotype switching (i.e. serum levels of IgM are at or above normal but levels of other isotypes are variably diminished or absent), because in these individuals the inability to class switch has a devastating effect on the response to certain pathogens and can increase the incidence of some autoimmune diseases [Durandy, 2006; Durandy, 2005; Minegishi, 2000; Revy, 2000].

Class switch recombination is a region-specific recombination event that occurs between two IgH switch regions, tracts of repetitive sequence (~1-12 kb) with high palindrome content which are positioned upstream of each of the constant regions [Chaudhuri, 2004; Kenter, 2004; Li, 2003; Neuberger, 2003; Ta, 2003; Tashiro, 2001; Zarrin, 2007], with the intervening DNA sequence looped-out and excised (i.e. deletional recombination [Barreto, 2003]) [Durandy, 2006; Durandy, 2005; Honjo, 2004; Lee, 2001; Muramatsu, 2000; Okazaki, 2002; Papavasiliou, 2002; Pham, 2005; Shinkura, 2004; Yu, 2005; Zarrin, 2004]. Deletion of specific switch sequences abrogates CSR to those constant region exons [Kenter, 2004; Zarrin, 2007]. Chromatin remodeling and sterile transcription from cryptic promoter sites upstream of switch regions, which contain sequence motifs that are recognized AID targets in SHM hotspots (i.e. RGYW) [Barreto, 2003; Zarrin, 2004], provide required access [Chaudhuri, 2003; Nambu, 2003] for AID and the recombination machinery [Durandy, 2006; Kenter, 2004; Lee, 2001; McBride, 2004; Okazaki, 2002; Yu, 2005; Zarrin, 2007]. As with

SHM, CSR efficiency has been correlated with transcription at switch regions [Lee, 2001; Ramiro, 2003], which also accrue mutations [Lee, 2001; Nagaoka, 2002; Okazaki, 2002; Yoshikawa, 2002; Zarrin, 2004]. The enhancer elements of the locus are of variable importance to CSR [Magor, 1999; Magor, 1994]. Activation of cryptic promoters for specific constant regions is influenced by the cytokines that are being produced as part of the immune response [Durandy, 2006; Durandy, 2005; Imai, 2003; Lee, 2001; Muramatsu, 2000; Nambu, 2003; Okazaki, 2002; Pham, 2005]; thus, the type of challenge (i.e. pathogen) the immune system is responding to further instructs it in how to tailor its effector response (i.e. through choice of functional antibody isotype).

Studies of the regions targeted during CSR have revealed that, in addition to the presence of repetitive AID target motifs (RGYW/WRCY), particularly the palindrome AGCT [Kenter, 2004; Ohta, 2006; Rada, 2002], these areas may form transient secondary structures, heteroduplexes known as R-loops, during transcription in an orientation-dependent manner [Bransteitter, 2006; Cascalho, 2004; Chaudhuri, 2004; Chaudhuri, 2003; Durandy, 2006; Durandy, 2005; Imai, 2003; Kenter, 2004; Lee, 2001; Nagaoka, 2002; Papavasiliou, 2002; Pham, 2005; Tashiro, 2001; Yu, 2005; Zarrin, 2004]. R-loops, which form when G-rich RNA transcript hybridizes with its DNA template, are accompanied by negative supercoiling of the DNA such that the ‘free’ non-template G-rich DNA strand coils around the RNA-DNA hybrid creating what is essentially an enlarged transcription bubble [Durandy, 2006; Durandy, 2005; Kenter, 2004; Ramiro, 2003; Yu, 2005]. Although a variety of complex DNA secondary structures, such as R-loops and stem-loops, are involved in cellular regulation and there are proteins that are implicated in the recognition of different conformations of higher-order structures of DNA [Ta, 2003; Tashiro, 2001], it is likely that the single-stranded DNA (ssDNA) generated in R-loops is more important to CSR than the R-loop structure itself [Kenter, 2004; Yu, 2005; Zarrin, 2004]. This notion is supported by the fact that R-loops are not a feature of SHM in the VDJ region [Chaudhuri, 2004], where mutability does not appear to be linked to secondary structures [Michael, 2002], and by data that show that the G-rich strand in switch

regions is dispensable for CSR in mice (i.e. can be replaced with AT-rich sequence from *Xenopus*) but its presence, correct orientation, AID target motif content, and length do improve efficiency [Kenter, 2004; Pham, 2005; Tashiro, 2001; Yu, 2005; Zarrin, 2004].

Similar to VDJ recombination, which utilizes a specially co-evolved system of enzyme and sequence to initiate recombination to co-opt elements of the DNA repair system of non-homologous end-joining (NHEJ) to recombine the DNA to form a complete variable region exon, CSR appears to have subverted NHEJ to recombine the DNA to replace IgH constant region exons when dsDNA breaks arise in switch regions during repair of AID-mediated lesions (discussed below) [Bransteitter, 2006; Lee, 2001; Li, 2003; Nagaoka, 2002; Okazaki, 2002; Wabl, 1999; Yu, 2005; Zarrin, 2007]. Indeed, that AID and switch regions are dispensable for CSR when replaced with a yeast endonuclease and its recognition site [Zarrin, 2007] suggests that the main purpose of AID in CSR is to generate excess repair to create dsDNA breaks as targets for NHEJ. Proteins for NHEJ (heterodimeric Ku70/Ku80 complexed with DNA-PKcs) and phosphorylated histone H2AX (γ-H2AX) foci, which form at dsDNA breaks [Brar, 2004; Papavasiliou, 2002; Woo, 2003], are involved in CSR but not in SHM [Barreto, 2003; Barreto, 2005; Durandy, 2005; Honjo, 2004; Li, 2003; Pham, 2005; Rada, 2002; Revy, 2000; Shinkura, 2004; Zarrin, 2007]. However, while it appears that NHEJ is the main path to CSR, it may not be the only path, as it appears that deficiency in some of the NHEJ molecules does not necessarily abrogate CSR [Durandy, 2005; Rada, 2002; Yu, 2005].

Refinement of antibody effector function by CSR mirrors the requirements and risks of diversification of antibody affinity by SHM. Although the outcomes of CSR and SHM have been studied for many years and both are known to make significant contributions to antibody affinity maturation, it was only recently discovered that a crucial mechanistic commonality links the initiation of the two processes: reliance on the newly identified activation-induced cytidine deaminase enzyme [Muramatsu, 1999].

1.2.7 Activation-Induced Cytidine Deaminase Initiates SHM And CSR

Expression of the mutator enzyme Activation-Induced Cytidine Deaminase (Figure 4) [Muramatsu, 1999; Muto, 2000], an approximately 24 kDa protein encoded by the Aicda gene [Muramatsu, 1999], in germinal center B-cells has been found to be essential for SHM [Muramatsu, 2000; Yoshikawa, 2002], CSR [Muramatsu, 2000; Nagaoka, 2002; Okazaki, 2002; Revy, 2000], and gene conversion [Arakawa, 2002; Harris, 2002] of the Ig genes [Durandy, 2006; Honjo, 2004; Neuberger, 2003; Ta, 2003]. The AID protein sequences translated from cDNA from both mouse and human have 198 amino acids and, while the two sequences do differ slightly (92% amino acid identity [Muto, 2000]), they exhibit complete conservation of sequence within their cytidine deaminase motifs [Muto, 2000]. Analyses of AID mutants suggest that this high level of evolutionary conservation is due to constraints that impact cytidine deaminase function and co-factor interactions [Durandy, 2006; Revy, 2000; Ta, 2003].

AID is related to the RNA-editing enzyme apolipoprotein B mRNA-editing catalytic subunit 1 (APOBEC1; 34% amino acid identity in mice [Muramatsu, 1999]), which converts mRNA that encodes apolipoprotein B (apoB) 100 into mRNA that encodes apoB48 by deaminating a specific cytidine residue (nucleotide 6666) in the former to introduce a stop codon (i.e. CAA to UAA) [Cascalho, 2004; Honjo, 2004; Ito, 2004; Jacobs, 2001; Maizels, 2000; Muramatsu, 1999; Neuberger, 2003; Petersen-Mahrt, 2002; Pham, 2005; Stavnezer, 2002; Ta, 2003], and to other APOBEC family members (e.g. APOBEC3G [Barreto, 2003; Cascalho, 2004; Kinoshita, 2006; Neuberger, 2005; Pham, 2005] and APOBEC3F [Pham, 2005] — which are involved in innate antiviral defense responses — and APOBEC2), and, even more distantly, to other cytosine deaminases [Conticello, 2005; Harris, 2002; Muto, 2000; Neuberger, 2003]. Based on preliminary phylogenetic analysis and predicted structural similarities, it was initially postulated that AID might accomplish SHM and CSR indirectly as an RNA-editing enzyme, similar in function to APOBEC1, that would edit a specific mRNA such that it would then encode a different

product — an endonuclease [Nagaoka, 2002; Ramiro, 2004] — required for SHM and/or CSR [Maizels, 2000; Muramatsu, 1999] or directly as a DNA cytidine deaminase [Cascalho, 2004; Durandy, 2005; Harris, 2002; Honjo, 2004; McBride, 2004; Neuberger, 2003; Petersen-Mahrt, 2002; Rada, 2002; Ramiro, 2004; Sohail, 2003; Stavnezer, 2002; Ta, 2003; Yu, 2005].

After the initial identification of AID, a variety of early observations and experiments in diverse systems with wild-type and mutant AID (e.g. constitutively hypermutating B-cell lines such as the centroblast-like Ramos Burkitt's lymphoma [Martin, 2002; Rada, 2002; Zhang, 2001], constitutive gene conversion B-cell lines such as DT40 chicken B-cell lymphoma [Arakawa, 2002; Di Noia, 2002; Harris, 2002], ectopic expression of AID in fibroblasts [Yoshikawa, 2002], pre-B-cell line 70Z/3 [Wang, 2004], hybridomas – approximately "plasma-like cells" [Martin, 2002], bacteria [E. coli] [Barreto, 2003; Li, 2003; Petersen-Mahrt, 2002; Ramiro, 2003; Sohail, 2003; Stavnezer, 2002; Ta, 2003], and yeast [Poltoratsky, 2004]) confirmed that its expression with an intact active-site correlated not only with the occurrence of hypermutation and gene conversion in DNA, but also with the recombination of DNA in systems that support recombination (e.g. CSR in the endogenous IgH locus of B lymphoma cells [Muramatsu, 2000] and splenic B-cells [Nagaoka, 2002], ectopic expression of AID in B-cells, T-cells, and fibroblasts with artificial switch constructs [Okazaki, 2002], and yeast [Poltoratsky, 2004]) suggesting that there are no other necessary factors as it is unlikely that they would be ubiquitously expressed among all of the very different experimental organisms [Durandy, 2006; Martin, 2002; Neuberger, 2003; Pham, 2005; Sohail, 2003; Yoshikawa, 2002]. It was also observed that certain AID deficiencies resulted in the alteration and/or abrogation of SHM [Honjo, 2004; Muramatsu, 2000; Rada, 2002] and/or CSR (e.g. deficiencies in AID are the cause of hyper-IgM syndrome type 2 [HIGM2] in humans [Durandy, 2006; Honjo, 2004; Li, 2003; Minegishi, 2000; Revy, 2000], AID knockout mice [Barreto, 2003; Muramatsu, 2000; Pham, 2005], mutant AID mice [Barreto, 2003; Shinkura, 2004]) and/or gene conversion (e.g. DT40 AID-/-

[Barreto, 2003; Harris, 2002]) and thereby impacted selection outcomes (e.g. repertoire usage) [Meffre, 2001].

Moreover, the expression levels of AID have an impact on the magnitude of these processes (i.e. dose-dependent) [Arakawa, 2002; Barreto, 2005; Martin, 2002; McBride, 2006; Okazaki, 2002; Rada, 2002; Wang, 2004; Yoshikawa, 2002; Zhang, 2001]. In fact, it appears that the level of AID expression must exceed an undefined threshold before mutations are seen [Michael, 2003; Zhang, 2001], beyond which the level of AID expression (measured as AID mRNA or AID-carrying virus/infection) was found to correlate with the amount of SHM (i.e. proportion of mutated cells, maximum number of mutations per cell, rate of SHM, and mutation frequency) [Martin, 2002; Rada, 2002; Yoshikawa, 2002; Zhang, 2001] or CSR [Okazaki, 2002] in hypermutating cell lines and when AID is ectopically expressed [Honjo, 2004]. That the amount of SHM is closely linked to the expression level of AID suggests that AID may be a limiting factor in the process [Barreto, 2005].

That hypermutation of DNA occurs when AID is ectopically expressed in bacteria [Harris, 2002; Petersen-Mahrt, 2002; Ramiro, 2003; Sohail, 2003] and yeast cells [Poltoratsky, 2004], organisms which would not be expected to encode the proposed target mRNA that is central to the hypothesis that AID may act indirectly as an APOBEC-like RNA mutator, strongly supports the hypothesis that AID initiates mutation via direct DNA deamination (Figure 4) [Honjo, 2004; Li, 2003; Neuberger, 2003; Ramiro, 2004]. Further support for the direct DNA deamination hypothesis is provided by the fact that AID is able to deaminate ssDNA *in vitro* [Bransteitter, 2003; Dickerson, 2003; Pham, 2003; Ramiro, 2004; Sohail, 2003; Ta, 2003]. But APOBEC1, a known RNA editor under physiological conditions, can cause a significant mutator phenotype by acting directly on DNA when overexpressed in bacteria (C:G to T:A transitions) [Bransteitter, 2003; Harris, 2002; Honjo, 2004; Pham, 2005; Ramiro, 2003] and *in vitro* on ssDNA [Durandy, 2005; Neuberger, 2003; Sohail, 2003]. In mammalian cells, APOBEC1 overexpression does not appear to lead to SHM or CSR [Kinoshita, 2006; Ta, 2003] even though it is associated with the development of cancer in certain

tissues [Harris, 2002; Pham, 2005]. Thus, AID's function as a deaminase of either RNA or DNA could not be conclusively established for some few years [Bransteitter, 2006; Imai, 2003; Kavli, 2007; Kenter, 2004; Maizels, 2000; Okazaki, 2002; Ramiro, 2004; Stavnezer, 2002], an undertaking that was further hampered by sometimes contradictory data from early AID experiments, due in part to technical difficulties (e.g. the AID protein proved difficult to purify for *in vitro* experiments) [Dickerson, 2003; Kinoshita, 2006; Larijani, 2005; Pham, 2005], differences in recombinant AID molecules that may have impacted deaminase activity or substrate interactions (e.g. fusion tags, phosphorylation state) [Brar, 2004; Dickerson, 2003; Harris, 2002; Yu, 2004; Yu, 2005], structural modeling disagreements [Pham, 2005], and difficulties identifying it in the nucleus where a direct DNA deaminator would necessarily need to be [Rada, 2002] which confounded attempts to analyze its function more directly [Besmer, 2004; Honjo, 2004; Kenter, 2004; Yu, 2004].

Subsequent experimentation has demonstrated that AID is, in fact, a DNA modifier [Chaudhuri, 2003; Dickerson, 2003; Kavli, 2005; McBride, 2004] that acts on ssDNA but not dsDNA [Bransteitter, 2003; Dickerson, 2003; Sohail, 2003] unless it is transcribed [Bransteitter, 2004; Bransteitter, 2006; Chaudhuri, 2003; Pham, 2005; Pham, 2003; Ramiro, 2003] to generate a suitable length of ssDNA target — *in vitro*, the empirically determined optimal appears to be ~26 nt, although "bubbles" as small as 3 to 5 nt are also targets [Bransteitter, 2003; Sohail, 2003; Yu, 2004]. Binding of AID to ssDNA does not rely on the presence of C despite this being the substrate of the enzyme [Bransteitter, 2004; Pham, 2005]. It appears that AID is able to deaminate C in the ssDNA of short switch region repeat R-loops *in vitro* [Yu, 2005] and, in B-cells in which CSR is occurring, AID has been found to be associated with the DNA of the IgH switch region [Besmer, 2004; Chaudhuri, 2004; McBride, 2004; Nambu, 2003]. But it is unknown precisely how AID is specifically targeted to the Ig gene for either SHM or CSR [Barreto, 2005; Cascalho, 2004; Durandy, 2006; Durandy, 2005; Kenter, 2004; Li, 2003; Neuberger, 2003; Ta, 2003; Yang, 2006]. Under normal physiological conditions, the catalytic APOBEC1 associates with a factor,

APOBEC1 Complementation Factor (ACF), that bears three binding motifs for mRNA and it has been postulated that AID may utilize a similar strategy of co-factor interaction for specific targeting to the Ig genes [Brar, 2004; Cascalho, 2004; Chaudhuri, 2004; Durandy, 2005; Harris, 2002; Honjo, 2004; Jacobs, 2001; McBride, 2006; Neuberger, 2003; Stavnezer, 2002; Yu, 2005]. Alternatively, it may rely on distinct structural elements of the target gene or its associated transcript [Barreto, 2005].

Although deaminase activity on free nucleotide substrates (dCTP) has been reported to occur in AID assays *in vitro* [Bransteitter, 2003; Muramatsu, 1999; Stavnezer, 2002], it appears that this activity may be due to low-level contamination of the recombinant AID preparations with cytidine deaminases derived from the expressor cells (e.g. E. coli) [Dickerson, 2003]. While AID does not bind to or act on the AU-rich mRNA target of APOBEC1 *in vitro* [Muramatsu, 1999], some recombinant AID preparations (i.e. from insect cells) have been found to bind to RNA *in vitro* [Bransteitter, 2006]; however, while these preparations required prior RNase treatment for deaminase activity to occur on ssDNA, the RNA interaction appears to be non-specific as AID does not act on the bound RNA [Bransteitter, 2003; Dickerson, 2003] and mutations in the active site that abrogate ssDNA binding do not appear to equally disrupt RNA binding [Dickerson, 2003]. That other experiments have not found an association of RNA with AID suggests that it typically does not bind RNA [Besmer, 2004; Kenter, 2004; Ta, 2003] but there is currently insufficient information to completely discount any biological relevance of the interaction [Bransteitter, 2006].

Specifically, under normal physiological conditions, the AID enzyme acts preferentially within RGYW/WRCY consensus sequences (discussed below) to directly deaminate cytosine to uracil in ssDNA (i.e. deoxycytidine to deoxyuridine) [Barreto, 2005; Chaudhuri, 2003; Dickerson, 2003; Pham, 2003; Sohail, 2003], which is actively generated from dsDNA through sterile transcription of the Ig genes [Bransteitter, 2006; Sohail, 2003] to which it is targeted somehow [Kenter, 2004] — possibly by interaction with RNA polymerase II [Nambu, 2003; Yang, 2006], replication protein A [Chaudhuri, 2004; Ramiro, 2004], or other

transcription elements [Barreto, 2005; Cascalho, 2004]. This targeted, enzyme-catalysed introduction of U into the Ig genes is a pivotal first step in the processes that lead to SHM and CSR outcomes [Cascalho, 2004; Neuberger, 2005; Petersen-Mahrt, 2002; Rada, 2002; Stavnezer, 2002].

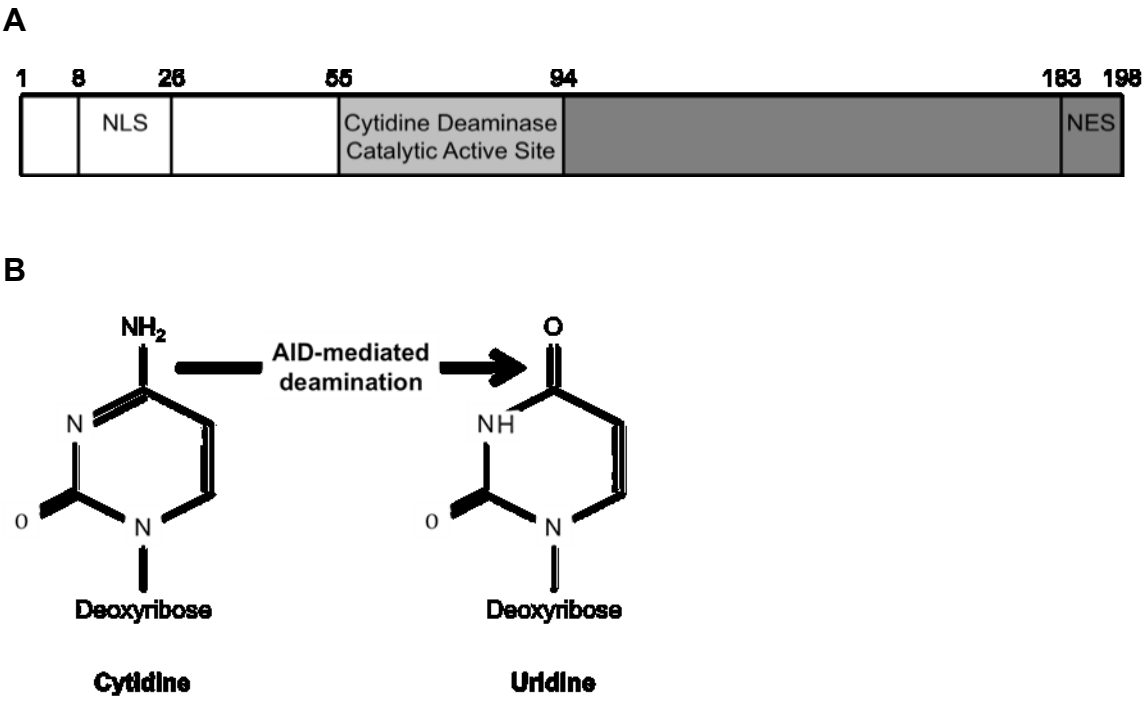


Figure 4. The mammalian activation-induced cytidine deaminase protein mediates deamination of cytosine to uracil in ssDNA.

A) The AID protein can be roughly divided into three domains based on function. The amino-terminal domain (white), which contains a putative bipartite Nuclear Localization Signal (NLS) sequence, is essential for somatic hypermutation. The middle domain, which contains the cytidine deaminase motif (light grey), is responsible for catalytic activity. The carboxy-terminal domain (dark grey), which contains a Nuclear Export Signal (NES) sequence, is required for class switch recombination. Numbers indicate amino acid residues for key features. B) AID catalyses the deamination of cytidine to uridine. Adapted from [Durandy, 2006; Durandy, 2005; Ichikawa, 2006; Pham, 2005; Ta, 2003].

1.2.7a AID-Mediated Deamination Initiates Repair

Although U is a standard component base of RNA, U in DNA is an aberration that often is a pre-mutagenic lesion [Kavli, 2007]. Lesions in DNA are detrimental to cells as they can lead to mutations or to blockages of transcription and/or replication, any of which may prove disruptive or lethal to the cell and, ultimately, the whole organism (e.g. oncogenic mutations) [Cascalho, 2004]. Uracil can arise in DNA (i.e. deoxyuridine) through non-enzymatic means (e.g. deamination of cytosine by alkaline hydrolysis or oxidative damage) and is only one of a multitude of different types of DNA lesions that are encountered by cells [Jacobs, 2001; Kavli, 2005; Kavli, 2007; Neuberger, 2003; Pham, 2005; Poltoratsky, 2007; Sohail, 2003; Stavnezer, 2002]. Consequently, a variety of mechanisms, many of which are somewhat redundant, have evolved that help cells to cope with and/or resolve potentially genotoxic DNA lesions through tolerance (e.g. translesion DNA polymerases) and/or repair (e.g. base excision repair) [Di Noia, 2002; Diaz, 1998; Kavli, 2005; Kavli, 2007; Neuberger, 2003; Pham, 2005; Stavnezer, 2002].

Intriguingly, despite the considerable effort exerted to maintain genomic stability, a variety of genome diversification mechanisms actively generate DNA lesions as intermediates [Arakawa, 2002; Bassing, 2002; Diaz, 2002; Papavasiliou, 2002; Ta, 2003]. These DNA lesions may recruit DNA damage tolerance and repair mechanisms, which typically act to stabilize the genome. However, in the context of diversification, the DNA damage tolerance and repair mechanisms are co-opted to generate genome alterations. The adaptive immune response utilizes this scheme for Ig gene diversification. Specifically, diversification of rearranged Ig genes relies on C deamination by AID to actively introduce pre-mutagenic U DNA lesions into the Ig genes in the variable region exons and switch regions. It is the replication or resolution of the AID-induced U DNA lesions by replicative DNA polymerases or the ubiquitously expressed DNA damage tolerance and repair mechanisms that determines the outcomes of SHM (e.g. mutational spectrum) in the Ig gene variable region and CSR (e.g. recombination) in the IgH gene constant region [Neuberger, 2003]. Petersen-

Mahrt et al. [Petersen-Mahrt, 2002] initially proposed a basic model outlining the means by which replicative DNA polymerases and DNA damage tolerance and repair mechanisms might accomplish the distinct SHM and CSR outcomes (discussed below and Figure 5). While the model has been essentially confirmed and further refined [Neuberger, 2005; Rada, 2004; Rada, 2002], it does not explain exactly how these normally very accurate processes are corrupted to generate the high rate of mutation seen in SHM instead of the usual correct repair. But data from gene expression studies that suggest that polymerase β, which does not appear to contribute to SHM [Diaz, 2002] and may even exert an anti-mutator effect [Poltoratsky, 2007; Poltoratsky, 2001], is downregulated in a hypermutating cell line [Poltoratsky, 2007] and that other more error-prone polymerases are upregulated [Poltoratsky, 2001] hint at programmed dysregulation of DNA damage tolerance and repair pathways in B-cells during SHM [Papavasiliou, 2002]. Indeed, activated B-cells exhibit preferential expression of particular repair molecules — uracil-DNA-glycosylase enzyme (UNG2) is upregulated while the partially redundant single-strand-selective monofunctional uracil-DNA-glycosylase enzyme (SMUG1) is downregulated [Kavli, 2007; Rada, 2004] — important in the repair of U lesions (discussed below) and may additionally regulate their function through specific patterns of post-translational modifications [Kavli, 2007]. Furthermore, it appears that this increase in error-proneness may be targeted to the Ig genes as there is a slight indication that repair in the Ig genes may be less efficient than in a housekeeping gene (dihydrofolate reductase) [Alrefai, 2007]. The AID enzyme itself, indispensable for generating the crucial U lesion, may play an additional role in directing or otherwise influencing the downstream response to the lesion, possibly through recruitment of specific DNA repair pathway molecules [Li, 2003], even though the introduction of the U lesion alone is sufficient to recruit repair [Stavnezer, 2002]. Thus, the chances of mutation occurring at or proximate to AID-mediated U DNA lesions are amplified as various elements conspire to both reduce repair fidelity and increase error-proneness at the sites of the Ig genes [Neuberger, 2005].

Under normal circumstances, repair of most U DNA lesions in the genome is accurate [Poltoratsky, 2007]. Accurate repair of DNA lesions is a multistep process that involves the recruitment and coordinated efforts of several ubiquitously expressed participants that act to recognize the type of DNA lesion, remove the damage from the DNA — a process that can itself generate a larger lesion [Cascalho, 2004] — and, ultimately, restore the integrity of the DNA. Uracil DNA lesions are commonly dealt with by the Base Excision Repair (BER) pathway wherein the U base is removed from the sugar of the DNA backbone by UNG [Jacobs, 2001; Kavli, 2005; Neuberger, 2003; Petersen-Mahrt, 2002; Poltoratsky, 2007; Rada, 2002; Stavnezer, 2002], which can recognize U in both ssDNA and dsDNA [Kavli, 2005; Kavli, 2007]. Excision of the resulting abasic site and end-processing are accomplished by apurinic/apyrimidinic endonuclease (APE), which nicks the DNA strand upstream of the abasic site to create a single-strand break in the DNA with a 3'-hydroxyl (3'-OH), and 5'-deoxyribose phosphate (5'-dRP) lyase (an activity intrinsic to mammalian polymerase β), which removes the 5'-dRP and leaves a ligatable 5'-phosphate end [Jacobs, 2001; Kovtun, 2007; Poltoratsky, 2007; Stavnezer, 2002]. The gap is filled in by a polymerase, such as the relatively accurate polymerase β, and then DNA ligase seals the nick to restore the integrity of the strand [Jacobs, 2001; Kovtun, 2007; Poltoratsky, 2007; Stavnezer, 2002].

But BER can be interrupted and/or superseded by tolerance mechanisms, which may be more lax — especially, as noted above, in B-cells undergoing SHM. Should replication proceed through the lesion before the repair process is initiated or completed, the mutation may be secured in the gene [Michael, 2002; Petersen-Mahrt, 2002; Rada, 2004]. Replicative DNA polymerases (e.g. polymerase ε and polymerase δ) treat U as though it is T (i.e. incorporate A opposite U). Thus, replication before repair does not result in stalled replication forks but does generate mutations (i.e. C:G to T:A). Unlike U, replicative DNA polymerases cannot cope with abasic sites in the template. These non-templating repair intermediates are lesions that result in the stalling of replication

forks. Stalled replication at abasic sites leads to the recruitment of translesion DNA polymerases (e.g. polymerase ι [Bransteitter, 2003; Neuberger, 2005; Poltoratsky, 2007; Poltoratsky, 2001], polymerase ζ [Diaz, 1998; Diaz, 2001; Durandy, 2006; Durandy, 2005], and polymerase θ [Bransteitter, 2006; Durandy, 2006]). Translesion DNA polymerases are DNA polymerases of lower fidelity than replicative DNA polymerases that enable a DNA lesion to be by-passed (i.e. tolerated) during replication rather than risk a permanent replication block by stalling high-fidelity replication to allow for repair [Diaz, 2002; Neuberger, 2005]. The prevalence of translesion DNA polymerases, also known as error-prone polymerases [Jacobs, 2001; Poltoratsky, 2001; Stavnezer, 2002], reflects their importance to the survival of cells but the trade-off for this increased replication efficiency and mitigation of immediate risks from blocked replication is an increase in mutations and the potential risks associated with them [Cascalho, 2004; Diaz, 2002; Poltoratsky, 2000]. Indeed, as these activated B-cells are rapidly proliferating, they may already be at increased risk for fixation of AID-mediated mutations through replication and translesion synthesis, even without concerted efforts to exploit the DNA damage tolerance and repair machinery to promote SHM outcomes.

If replication does proceed past the U before any repair occurs, it will typically serve as template for incorporation of an A in the nascent strand, instead of the G that should be incorporated, which will yield a daughter cell with a U:A base pair [Stavnezer, 2002]. Subsequent repair of the U, or further replication from the A-containing DNA strand, will result in a C to T transition at the initial site of AID action (phase one A mutation) [Di Noia, 2002; Neuberger, 2005; Neuberger, 2003; Petersen-Mahrt, 2002]. Although the phase one A mutations of SHM are typically restricted to “transitions only” in the original model [Petersen-Mahrt, 2002], more recent SHM studies with various polymerase-deficient mice have indicated that U may also occasionally serve as template for the incorporation of bases other than A by some translesion DNA polymerases such as Rev1, which

instead incorporates C to create C to G transversions at the site of the AID-mediated U lesion [Bransteitter, 2006].

If the repair process is initiated but replication does proceed before it is complete, then transition or transversion mutations (phase one B mutation) may occur [Di Noia, 2002; Neuberger, 2005; Neuberger, 2003; Petersen-Mahrt, 2002]. Likewise, if repair gets completed, but is inaccurate, then either transition or transversion mutations may result.

In addition to BER led by UNG — which can also interact with pathways other than BER [Kovtun, 2007; Neuberger, 2005] — repair may proceed via other repair pathways (e.g. patch repair) that yield transitions or transversions (phase two mutations) [Di Noia, 2002; Neuberger, 2005; Petersen-Mahrt, 2002; Rada, 2002]. This is likely to be the cause of mutation at A:T sites proximate to the initial U lesion as these mutations appear to require MSH2/MSH6 and DNA polymerase η [Aoufouchi, 2008; Barreto, 2005; Bransteitter, 2006; Diaz, 2002; Durandy, 2005; Kavli, 2007; Neuberger, 2005; Neuberger, 2003; Papavasiliou, 2002; Pham, 2005; Rada, 2004; Rada, 2002; Rada, 2001; Rada, 2002; Rogozin, 2004; Stavnezer, 2002; Zhang, 2001].

More recently, the alternative MMR that contributes to SHM in the G1 phase of the cell cycle has been redefined as “non-canonical MMR” (ncMMR) to distinguish it from “canonical replication-associated MMR”, which is less error-prone [Pilzecker, 2019]. And it has been proposed that five pathways may operate to resolve pre-mutagenic U DNA lesions during SHM [Pilzecker, 2019]. It appears that several elements (e.g. UNG, MSH2, and pol η) may contribute to more than one pathway [Pilzecker, 2019]. The first three pathways generate mutation at C:G (i.e. phase one mutations) and the last two pathways generate mutation at A:T (i.e. phase two mutations) [Pilzecker, 2019]. In the first pathway, replication past U generates C:G to T:A transitions [Pilzecker, 2019]. In the second pathway, UNG and translesion synthesis generate transitions and transversions at C:G [Pilzecker, 2019]. In the third pathway, it appears that UNG and ncMMR cooperate to generate transitions and transversions at C:G, but the details of this mechanism are poorly understood [Pilzecker, 2019]. It appears that

pathways two and three may be equally responsible for the generation of transversions at C:G [Pilzecker, 2019]. In the fourth pathway, which is the primary pathway of A:T mutagenesis, ncMMR generates mutation at A:T [Pilzecker, 2019]. In the fifth pathway, mutation at A:T is dependent on UNG and ubiquitinated proliferating cell nuclear antigen [Pilzecker, 2019].

The scenario presented for SHM is similar to that of CSR wherein AID is instrumental in initiating the U DNA lesions in transcribed IgH gene switch regions [Chaudhuri, 2004; Chaudhuri, 2003; McBride, 2004; Nambu, 2003] that lead to CSR during repair [Cascalho, 2004; Nagaoka, 2002; Stavnezer, 2002]. The attempted resolution of clusters of AID-mediated U lesions on both DNA strands [Yu, 2004], by UNG led BER [Imai, 2003] or other mismatch repair pathways (e.g. MSH2/MSH6) [Rada, 2004; Rada, 2002] possibly coupled with replication [Poltoratsky, 2004], introduces a critical number of single-strand breaks to both DNA strands within a compact region of sequence leading to double-strand breakage which is resolved through a region-specific recombination event that requires NHEJ [Bransteitter, 2006; Durandy, 2006; Honjo, 2004; Jacobs, 2001; Nagaoka, 2002; Neuberger, 2003; Petersen-Mahrt, 2002; Pham, 2005; Rada, 2002]. These processes also give rise to point mutations and small deletions within the switch regions [Nagaoka, 2002], the latter of which are almost entirely absent in MSH2-/-UNG-/- double-knockout mice, which also have extremely limited CSR [Rada, 2004]. In the absence of AID, neither mutations nor phosphorylated histone H2AX (γ-H2AX) foci, which are associated with dsDNA breaks [Papavasiliou, 2002], are found to be associated with switch regions [Honjo, 2004; Nagaoka, 2002; Shinkura, 2004].

It appears that differential utilization of DNA damage tolerance and repair mechanisms contributes significantly to the different outcomes of AID-mediated U DNA lesions, with both CSR (e.g. NHEJ) and gene conversion (e.g. XRCC2 and XRCC3 and RAD51B [Arakawa, 2002; Barreto, 2003; Di Noia, 2002; Diaz, 2002; Honjo, 2004; Kenter, 2004; Neuberger, 2003; Papavasiliou, 2002; Petersen-Mahrt, 2002]) requiring repair molecules that are not needed for SHM [Durandy, 2006; Li, 2003; Nagaoka, 2002; Neuberger, 2003] and AT-focused

mutation requiring repair molecules (e.g. MSH2) that are not required for GC-focused mutations (e.g. UNG) [Rada, 2004; Rada, 2002]. Additionally, it seems that the reliance on replicative DNA polymerases and ubiquitous DNA damage tolerance and repair mechanisms for AID-mediate outcomes is the reason AID-mediated hypermutation can occur — through replication, failure of non-replicative repair [Poltoratsky, 2000], and/or translesion synthesis of AID-induced U lesions — even when AID is ectopically expressed in other cell types or organisms [Neuberger, 2005; Ramiro, 2003].

The expression of AID on UNG-deficient or inhibited (i.e. Ugi) backgrounds highlights the importance of UNG for normal outcomes in AID-mediated processes and provides further support for AID directly deaminating DNA [Di Noia, 2002; Imai, 2003; Li, 2003; Neuberger, 2003; Pham, 2005; Pham, 2003; Poltoratsky, 2004; Rada, 2004; Ramiro, 2004; Sohail, 2003; Stavnezer, 2002] as the resulting mutations, which differ from the spectrum of background mutations, are predominantly C to T transitions. Additionally, when AID is expressed in the absence of UNG in E.coli, there is a significant transcription-dependent increase in mutation frequency [Ramiro, 2003; Sohail, 2003] well above the sum of mutations seen when AID is expressed or UNG is deficient separately [Besmer, 2004; Di Noia, 2002; Petersen-Mahrt, 2002; Rogozin, 2004]. This suggests that UNG is the predominate initiator of the repair of U DNA lesions in these bacteria and, in its absence, the U DNA lesions introduced by AID may be more likely to be tolerated than dealt with by potential backup enzymes (e.g. Mug) [Petersen-Mahrt, 2002]. Intriguingly, increases in genomic U DNA lesions, due to dUTPase deficiency rather than AID transfection, are known to increase recombinations in UNG-proficient cells [Rada, 2002], supporting the CSR hypothesis that excess repair may be sufficient to trigger the recombination event.

When AID is expressed in a yeast system, mutations in haploids and mitotic recombination in diploids (i.e. both intrachromosomal and interchromosomal in a region that incorporates 13 sequences that are part of mouse Sμ repeats) can occur [Poltoratsky, 2004]. An UNG deficient background in this system leads to an increase in mutations, but a decrease in recombination [Poltoratsky, 2004].

The increase in mutations is consistent with the U DNA lesions being tolerated rather than repaired (i.e. increase in C to T and G to A transitions) [Poltoratsky, 2004; Rada, 2004], which supports the view that AID-mediated CSR is likely brought about by breaks introduced as part of the repair process. That the removal of U DNA damage, rather than tolerance, may be important to initiate a program of recombination instead of mutation is further corroborated by the fact that remediation of other forms of DNA damage (i.e. alkylation) by BER can likewise lead to recombination in bacteria, yeast, and mammals [Poltoratsky, 2004].

In human B-cells, expression of the nuclear isoform of UNG is increased during CSR [Imai, 2003] and the absence of UNG function (i.e. hyper-IgM syndrome type 5) leads to disturbances in SHM and inhibition of CSR [Cascalho, 2004; Durandy, 2005; Imai, 2003; Li, 2003; Neuberger, 2003; Rada, 2004; Rada, 2002; Shinkura, 2004]. Despite the fact that there are at least four other enzymes in mammals that may remove U from DNA [Imai, 2003; Kavli, 2005; Rada, 2002], defects specifically in UNG that preclude its ability to remove U lesions from DNA lead to significantly reduced CSR (i.e. HIGM) [Imai, 2003; Kavli, 2005; Pham, 2005; Rada, 2002], altered SHM metrics (e.g. mutation spectrum and frequency) [Honjo, 2004; Imai, 2003; Kavli, 2007; Pham, 2005; Rada, 2004; Rada, 2002], and an increase in the prevalence of U lesions throughout the genome [Kavli, 2005]. That SHM other than phase one A mutations and CSR aren't completely blocked in the absence of UNG function, especially in mice where there is only a slight increase in SHM, indicates that redundant repair mechanisms can supplement for it but to a much lesser degree [Bransteitter, 2006; Cascalho, 2004; Di Noia, 2002; Honjo, 2004; Imai, 2003; Rada, 2004; Rada, 2002]. Indeed, a study by Rada et al. [Rada, 2004] found that ectopic overexpression of SMUG1 in B-cells was not able to compensate for UNG deficiency [Kavli, 2005]. While this may be due in part to differences in activity levels and substrate preferences of the enzymes [Kavli, 2005], it has also been suggested that the apparent inability of other Uracil-DNA-glycosylase enzymes or repair pathways to adequately compensate for UNG deficiency in this context is because the role of

UNG is more complex than mere U removal [Honjo, 2004; Kavli, 2005; Rada, 2002]; that is, beyond its catalytic function, UNG may play a specific role downstream of AID action to recruit other molecules involved in SHM or CSR [Kavli, 2007]. The recruitment hypothesis is supported by the finding that CSR can be restored in UNG-deficient B-cells that are supplied with certain catalytically-inactive UNG mutants [Barreto, 2005; Rada, 2004].

Interestingly, it appears that, although UNG is necessary for the occurrence of normal SHM [Kavli, 2007], A:T mutation is not affected in UNG-knockout mice [Imai, 2003; Rada, 2002] and Rada et al. [Rada, 2004] propose that this might indicate that phase two mutations can be produced by MSH2/MSH6 heterodimer recognition of the U DNA lesion independent from UNG and/or BER, an idea supported by sequence data from MSH2-/- mice which shows that they lack AT-focused mutation.

In addition to UNG, deficiencies in other DNA damage tolerance and mismatch repair mechanisms (e.g. MSH2 [Rada, 2004] and MSH6 [Neuberger, 2005]) have been found to variably impact SHM metrics (e.g. mutation distribution) and capacity for CSR, which suggests that they play important roles in these processes [Barreto, 2003; Bransteitter, 2006; Di Noia, 2002; Jacobs, 2001; Kavli, 2005; Kavli, 2007; Kovtun, 2007; Li, 2003; Martin, 2002; Pham, 2005; Rada, 2004; Rada, 2002; Wang, 2004].

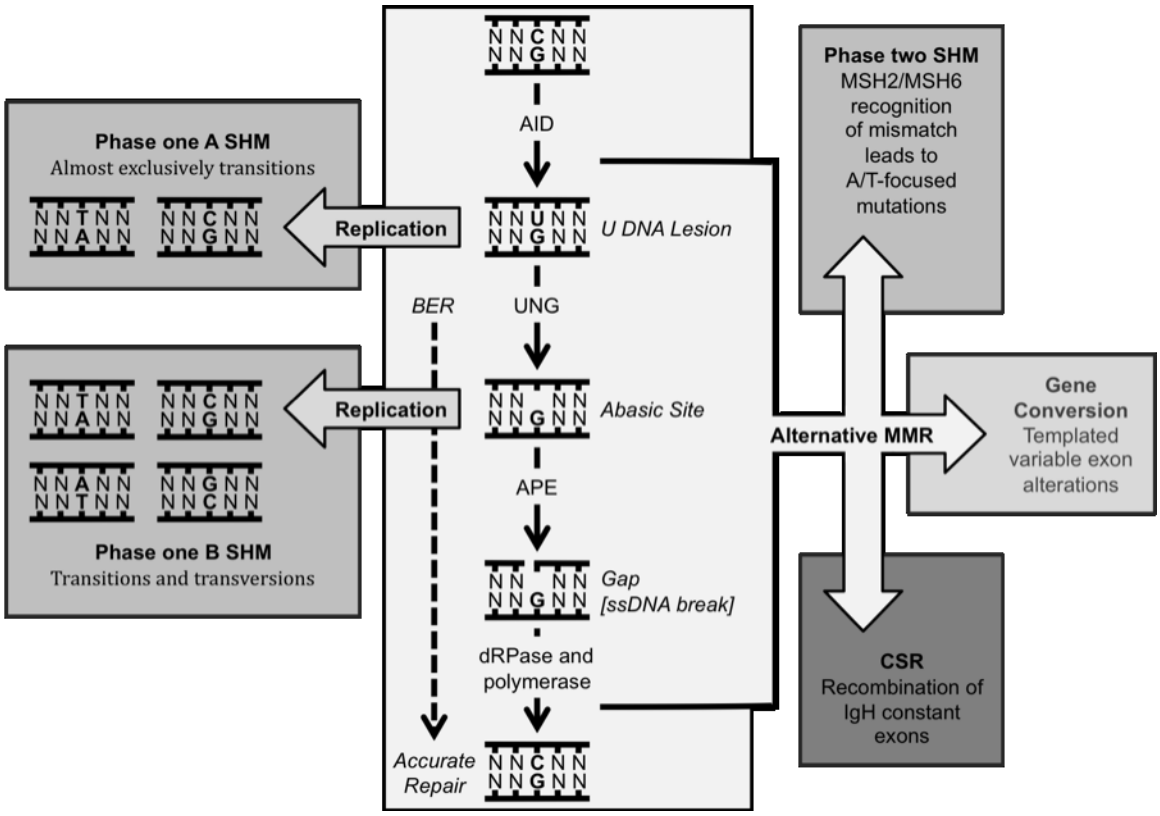


Figure 5. The AID-mediated deamination of cytidines within the immunoglobulin genes of centroblast B-cells engages DNA damage tolerance and repair mechanisms that promote SHM and CSR outcomes.

During antibody diversification, it appears that recognition and repair of AID-mediated pre-mutagenic U DNA lesions by the Base Excision Repair (BER) pathway (center box and detailed in the main text) may be interrupted or supplanted by various other DNA damage tolerance and repair mechanisms that are more mutagenic. The mechanisms that contribute to SHM outcomes can be roughly divided into three groups: Phase one A and B SHM generate mutations focused at C:G and Phase two SHM generates mutations that can also occur at A:T. Phase one A SHM mutations (upper left box), which are almost exclusively C:G to T:A transitions, result from the replication of U by replicative DNA polymerases, which are unable to distinguish U from T in DNA templates. At sites of U (i.e. rightly C), these polymerases incorporate A into the nascent DNA strand instead of G. The A then templates for T during additional replication or

repair of U. Phase one B SHM mutations (lower left box), which include both transitions and transversions at C:G, result from the interruption of repair by replication with error-prone translesion polymerases that bypass the non-templating abasic site. Phase two SHM mutations (upper right box) result from the resolution of the lesion by alternative mismatch repair mechanisms. Phase two SHM is not as restricted to the site of the original lesion and gives rise to mutations at A:T (i.e. responsible for WA hotspots). In CSR (lower right box), it appears that repair of clusters of pre-mutagenic U DNA lesions in switch regions generates a critical number of single-strand DNA breaks in both DNA strands that are resolved through recombination. Gene conversion (middle right box), which creates diversity through imprinting, has not been found to contribute to antibody diversification in the mouse and human. AID = Activation-Induced Cytidine Deaminase, APE = Apurinic/apyrimidinic Endonuclease, CSR = Class Switch Recombination, SHM = Somatic Hypermutation, UNG = Uracil-DNA Glycosylase. Adapted from [Bransteitter, 2006; Neuberger, 2005; Neuberger, 2003; Petersen-Mahrt, 2002; Poltoratsky, 2004; Rada, 2004; Rada, 2002].

1.2.7b AID Preferentially Deaminates C Within Consensus Sequences

Combined Ig and non-Ig gene sequence data obtained from a selection of *in vivo* and *in vitro* studies suggest that the AID enzyme does not recognize a specific primary DNA sequence, but preferentially targets C in a consensus RGYW/WRCY sequence with a detectable hierarchy of preferences (i.e. different deamination efficiencies) [Barreto, 2005; Bransteitter, 2003; Bransteitter, 2006; Li, 2003; Pham, 2005; Sohail, 2003; Ta, 2003; Yu, 2004]. It appears that, although AID remains deamination competent under an array of conditions [Dickerson, 2003], the particular experimental context can have an impact on targeting [Bransteitter, 2006] as AID preparations appear to have a sole *in vitro* preference for WRC [Pham, 2003; Rogozin, 2004; Wang, 2004], possibly with a slight preference for purine immediately downstream of the motif (WRCr) [Pham, 2003; Yu, 2004; Yu, 2005], *in vitro* reaction conditions may affect AID targeting [Sohail, 2003], and the strength of preference for hotspots over coldspots varies depending on accessibility of the target DNA (i.e. higher in transcribed dsDNA substrate than gapped ssDNA substrate) [Bransteitter, 2004]. And AID activity and target preference are also affected by the substrate as “bubble” DNA (5 or 9 nt) is targeted more than ssDNA [Bransteitter, 2003] and the target C preference of ssDNA is different from that of the ssDNA of R-loops generated in transcribed plasmids [Yu, 2005].

While the AID consensus target is consistent with the RGYW/WRCY SHM hotspot motif [Bransteitter, 2004; Li, 2003; Pham, 2005; Pham, 2003], AID does not act at WA/TW, which is also recognized as a mutational hotspot of Ig gene SHM [Pham, 2005; Rada, 2001], as these sites contain no C [Neuberger, 2005]. The significant contributions and variable application of the diverse replicative DNA polymerases and DNA damage tolerance and repair mechanisms to SHM outcomes explain not only why different mutations can occur at the immediate site of an AID-deaminated C, but also why mutations found at the sites of other bases do not invalidate the AID direct DNA deamination model of SHM [Bransteitter, 2004; Durandy, 2005; Neuberger, 2005; Rada, 2004]. That is,

despite not being substrates of AID, bases other than C are often sites of mutation in AID-mediated SHM but these apparently non-conforming mutations, including those in WA/TW sequence context [Rada, 2004; Rogozin, 2004], reflect the diverse functions, sequence preferences, biases, and fidelities of different DNA damage tolerance and repair molecules and pathways [Di Noia, 2002; Neuberger, 2005]; indeed, Ig genes may have evolved at the sequence level to additionally exploit these mechanisms at specific sites, as has already been noted for AID targeting, to further refine the mutational potential at SHM hotspots (WRC) and coldspots (SYC) [Bransteitter, 2004; Pham, 2005; Pham, 2003; Poltoratsky, 2001] and within their hierarchies [Besmer, 2004; Larijani, 2005; Milstein, 1998; Neuberger, 1995; Sohail, 2003; Zarrin, 2004]. Thus, mutations generated proximate to the site of the initial U DNA lesion may be the result of further resection of the DNA undergoing repair due to the removal of flanking DNA in addition to the damaged DNA and/or to overprocessing of the DNA as damage may be somewhat clustered in the region and/or repair enzymes may be in competition with transcription or replication machinery for access to the DNA [Jacobs, 2001]. Resection of the DNA during repair of AID-mediated U lesions may also be the reason that the frequency of mutation of the bases immediately 5' of the transcription start site is slightly above background despite this point being considered as the 5' boundary of SHM [Gordon, 2003; Rada, 2001].

The potential for overprocessing of the DNA during repair of AID-mediated U lesions, although most conspicuous in the densely clustered hotspots of the IgH switch regions, may be heightened in the VDJ exon by the apparent processive action of AID [Bransteitter, 2004; Bransteitter, 2006; Larijani, 2005; Pham, 2005; Pham, 2003]. Using an *in vitro* system, Pham et al. [Pham, 2003] found that, while only a small percentage of the ssDNA substrates (~2%) were targeted for deamination by AID, those that were targeted by AID tended to have several deaminated residues (>70% had 11-80). Increasing the incubation time of AID with the ssDNA substrate did have a slightly positive impact on both of these factors but only over a range of a few minutes (<10 minutes) [Bransteitter, 2004; Pham, 2005]. Interestingly, the proportion of deaminations within hotspots was

highest in the substrates with the fewest deaminations and diminished as deaminations increased (i.e. hotspots became saturated) [Bransteitter, 2004; Pham, 2005].

The influences of DNA damage tolerance and repair enzymes also likely explain some of the minor differences in the target consensus sequence identified in mutation data obtained from *in vitro* versus *in vivo* hypermutation experiments in the absence of selection. This is supported by a review of select sequence data by Rogozin and Diaz [Rogozin, 2004] who found that DGYW/WRCH (D=A,G,T and H=T,C,A) may be a “better predictor” of mutation sites in Ig and non-Ig genes *in vivo* in B-cells and fibroblasts. They speculate that this may be due to more efficient repair of mutations that occur at CGYW/WRCG sites, possibly by those elements that have evolved to maintain the integrity of CpG islands in the genome, rather than to actual differences in AID targeting [Larijani, 2005; Pham, 2003; Rogozin, 2004].

Methylation at CpG sites may also impact SHM more directly. Early research *in vitro* indicated that methylation of C in ssDNA substrates confers variable protection against deamination by AID specifically at the site of the 5-methylcytosine (5-mC) residue (i.e. *in vitro* deamination efficiency at 5-mC appears to be 10% - 30% of that at C) with only minor influence, if any, on deamination at nearby C residues and both effects appeared to be affected by sequence context [Bransteitter, 2003; Larijani, 2005; Pham, 2005]. That methylation of C may be able to exert minor influence on AID targeting, suggests that methylation at CpG in non-Ig genes may be slightly protective during AID-mediated antibody diversification [Bochtler, 2017; Bransteitter, 2003; Larijani, 2005; Pham, 2005]. Whether methylation might be relevant as a control mechanism in the physiological context, where it would be coupled with other mechanisms that regulate AID, is unclear [Larijani, 2005; Pham, 2005]. It is unlikely to contribute to mutational control in highly expressed genes (e.g. Ig), as these genes have little or no methylation, or in repressed genes, despite these genes being heavily methylated, as AID cannot access these genes (i.e. no ssDNA). But, in some moderately expressed non-Ig genes, even the apparently

small amount of protection from or redirection of AID-mediated deamination that might be provided by 5-mC could be important.

More recently, a role has been proposed for AID in the demethylation of 5-mC *in vivo* during epigenetic reprogramming (see Discussion) [Bochtler, 2017; Moris, 2014; Villota-Herdoiza, 2013]. While it appears to be the deaminase activity of AID that leads to demethylation of 5-mC, the precise mechanism is still under investigation [Bochtler, 2017; Moris, 2014]. Thus, although AID deamination of 5-mC does appear to be less efficient than C, that AID has been implicated in the targeting of 5-mC for demethylation *in vivo* indicates that methylation of C likely does not provide much protection from off-target mutation as a result of AID activity [Bochtler, 2017].

Repair preferences may also partly contribute to another perceived difference in AID-mediated mutation observed among different organisms *in vivo* and *in vitro*. Both coding and noncoding strands of the Ig genes are recognized as targets of SHM [Milstein, 1998; Ramiro, 2004] but there has been some confusion and controversy over whether SHM exhibits strand bias (polarity) [Diaz, 2001; Green, 1998; Jacobs, 2001; Michael, 2002; Milstein, 1998; Neuberger, 1995; Peters, 1996; Poltoratsky, 2000]. Expression studies with mammalian AID *in vitro* [Bransteitter, 2004; Chaudhuri, 2004; Pham, 2003; Zarrin, 2004] and in bacteria [Yu, 2004] suggest a real bias for AID-mediated mutation of the non-template strand (i.e. coding strand) [Barreto, 2005; Bransteitter, 2006; Ramiro, 2003; Sohail, 2003], which would be exposed on the prokaryotic RNA polymerase during transcription [Chaudhuri, 2004; Ramiro, 2004]. But this does not seem to be the case in yeast [Poltoratsky, 2004] or in mammalian B-cells [Bransteitter, 2006; Ramiro, 2004] where current data appear to support an absence of strand bias in AID activity [Brar, 2004; Kavli, 2005; Li, 2003; Rada, 2002], possibly as a consequence of differences in DNA accessibility during transcription by eukaryotic RNA polymerase II or the interaction of modifying co-factors [Chaudhuri, 2004]. It has also been suggested that the transcribed strand may be made more accessible to AID during repair of the non-transcribed strand (i.e. similar to a “gapped” substrate) [Pham, 2005].

Confoundingly, although the AID enzyme does not seem to exhibit strand bias in mammalian B-cells, there is an apparent strand bias in the mutations in AID-expressing cells [Barreto, 2005; Milstein, 1998; Rada, 2002; Rogozin, 2004]. Indeed, differences have been observed in the mutation frequencies in the coding strand, where purines (A and G) are mutated more frequently than pyrimidines (C and T) [Diaz, 1998; Milstein, 1998]. That this is likely due to repair biases downstream of AID action is most clearly supported by mutation at sites of A:T base pairs (i.e. nucleotides that are not direct targets of the AID enzyme) where A is mutated more often than T in the coding strand [Barreto, 2005; Bransteitter, 2006; Jacobs, 2001; Neuberger, 2005; Neuberger, 1995; Rada, 2002] — mutations which are abrogated in MSH2-/- mice [Rada, 2004; Rada, 2002]. Sequence bias may play a role [Milstein, 1998] and transcription-coupled repair, in which the transcribed (i.e. template) strand is preferentially repaired [Alrefai, 2007], may also contribute to give the appearance of strand bias [Green, 1998; Jacobs, 2001; Neuberger, 2005]. Thus, repair may not only obscure the actual site of AID action in SHM, it may also give the appearance of strand bias [Jacobs, 2001; Ramiro, 2003].

While the actions of various DNA damage tolerance and repair mechanisms have been identified as the source of some of the differences observed between physiological SHM in the Ig genes of germinal center B-cells and AID-mediated outcomes under various experimental conditions, the precise reasons for some other differences are less clear. For instance, although the AID enzyme reproduces several features of physiological Ig gene SHM, such as a blunt 5' boundary and a diffuse 3' boundary, on transcribed dsDNA substrates *in vitro*, the mutations mostly occur over a significantly shorter distance (a few hundred bp versus ~ 2 kb) and begin within a few bases downstream of the promoter rather than approximately 200 bp downstream [Bransteitter, 2004; Bransteitter, 2006]. Among AID substrates *in vitro*, mutations are spread more uniformly over ssDNA (e.g. gapped) substrates than they are in transcribed dsDNA substrates [Bransteitter, 2004]. Intriguingly, promoter-proximal mutations and a more uniform spread of mutations are also observed in the Ig variable region of BL2

cells that have been stimulated and treated with a histone deacetylase inhibitor or that ectopically overexpress AID without stimulation [Woo, 2003].

Ectopic and/or overexpression of AID can lead to global genome hypermutation [Okazaki, 2003; Wang, 2004], the outcome of which can be tumorigenesis [Cascalho, 2004; McBride, 2006; Okazaki, 2003; Pham, 2005] or autoimmune disorders [Wedemayer, 1997] and shortened life-span [Honjo, 2004; Kinoshita, 2006; Okazaki, 2003]. This reinforces that hypermutation does not require Ig gene sequences to occur but they can have an impact on targeting [Wang, 2004]. Directed evolution (i.e. for enhanced fluorescence characteristics) of a target fluorescent protein gene randomly integrated into a population of AID-expressing cells indicated that, despite integration throughout the genomes of the initial cell population and the occurrence of mutation at these diverse locations, the evolved population contained only one integration site at the IgH gene locus [Wang, 2004]. While this suggests that proximity to the endogenous IgH gene sequence enhanced mutability of the exogenous target, it is possible that other factors related to the integration site affected the mutability of the transgene [Aoufouchi, 2008; Wang, 2004]. And, just as some non-Ig genes are more frequent sites of SHM, some non-Ig genes are less frequently mutated than other non-Ig genes [Michael, 2003; Wang, 2004], including some that are highly transcribed [Larijani, 2005; Woo, 2003]. This differential mutability of genes may partly explain why ubiquitous constitutive expression of AID in mice appears to generate tumors in only a very few cell types (i.e. mainly T-cells and lung epithelia) [Cascalho, 2004; Durandy, 2005; Honjo, 2004; Kinoshita, 2006; Okazaki, 2003], although other factors, such as cell turnover rates, likely impact susceptibility to tumor production.

Interestingly, Gramlich et al. [Gramlich, 2012] physically measured the proximity of a selection of Ig and non-Ig genes, including some that are mutated by AID (i.e. off-target mutation) and some that are not, in activated B-cells. Transcriptionally active genes are known to dynamically relocate to shared transcription foci [Osborne, 2004]. This transient spatial reorganization to preassembled transcription “factories” can bring otherwise physically distant

genes — both cis and trans — into close proximity [Osborne, 2004]. The c-Myc gene, which is prone to AID-induced translocation to the IgH gene, has been found to preferentially relocate to the same transcriptional factory as the IgH gene in stimulated B-cells [Gramlich, 2012]. This suggests that proximity to the physiological targets of AID (i.e. Ig genes) may be a contributing factor in the off-target activity of AID [Gramlich, 2012]. But the investigation by Gramlich et al. [Gramlich, 2012] did not identify a correlation between off-target AID mutation of non-Ig genes and their proximity to Ig genes during AID expression. Moreover, no preferential association was found between hypermutating IgL and IgH genes. However, measurements were obtained from cells in interphase while AID activity has been found to occur during early G1 [Gramlich, 2012].

Ectopic expression of AID in non-B-cells, including non-mammalian cells, still leads to mutations that target RGYW motifs [Martin, 2002; Petersen-Mahrt, 2002; Ramiro, 2004] but these occur predominantly at GC [Aoufouchi, 2008; Brar, 2004; Kavli, 2007; Martin, 2002; Petersen-Mahrt, 2002; Rada, 2002; Ramiro, 2004; Ta, 2003; Yoshikawa, 2002], are C to T and G to A transitions [Aoufouchi, 2008; Petersen-Mahrt, 2002; Stavnezer, 2002; Yoshikawa, 2002], and exhibit distributions that differ from background mutations [Poltoratsky, 2004]. A similar mutational pattern has been observed in hypermutating cell lines such as 18-81 and Ramos [Bachl, 2001] [Martin, 2002; Rada, 2002]. In Ramos, the SHM appears to be unstable over time and prolonged culture generates non-mutating cells that express lower levels of AID [Zhang, 2001], possibly due to mutations accruing in the AID gene [Aoufouchi, 2008; Gordon, 2003; Kinoshita, 2006]. Overexpression of AID in B-cells also leads to increased whole genome mutation but the spectrum is more normal with mutations occurring at all four nucleotides [Aoufouchi, 2008; Ramiro, 2004; Wang, 2004], although there is a notable reduction in A:T mutation in B-cell lines [Barreto, 2005; Zhang, 2001]. Regardless of the cell type, the level of transcription of the mutated genes in AID-expressing cells is still central to their mutation frequency [Sohail, 2003; Wang, 2004; Yoshikawa, 2002]. More recently, it has been found that transcription from sites other than the Ig promoter and cryptic promoter sites ahead of switch regions

may contribute to the level of transcription in the Ig locus [Wang, 2017; Yeap, 2019]. That is, convergent transcription has been proposed to enhance AID activity, possibly by generating more access to ssDNA [Wang, 2017; Yeap, 2019]. Wang et al. [Wang, 2017] found that convergent transcription occurs at AID targets during early G1 (i.e. when AID-mediated C deamination occurs).

Exploration of the differences between legitimate physiological SHM in the Ig genes of germinal center B-cells and mutations generated by AID overexpression in other cell types, in the absence of other factors, or *in vitro* may help to reveal the important components of SHM [Bransteitter, 2006; Gordon, 2003]. For instance, activated B-cells may express factors that restrict AID activity to the Ig genes as, despite expressing an abundance of AID, mutation in these cells does not indiscriminately target transcribed non-Ig genes to the extent that it does when expressed in other cell types [Besmer, 2004; Durandy, 2005].

1.2.7c AID Protein Mediates SHM And CSR Via Functionally Distinct Regions

Precisely how such distinctly different ends as SHM and CSR are brought about by essentially the same means (e.g. AID, ubiquitous DNA damage tolerance and repair factors, and transcription) is unclear but it is likely that several layers of regulation contribute to each process [Cascalho, 2004; Nagaoka, 2002]. Data from studies of immune-deficient individuals and mutant AID or repair molecules suggest that different domains of AID are, in part, responsible for leading to different outcomes, but it is still not understood how these specific functions are uncoupled and independently regulated [Barreto, 2003; Cascalho, 2004; Durandy, 2006; Imai, 2003; Li, 2003; Poltoratsky, 2000; Shinkura, 2004; Ta, 2003].

For the purposes of description and discussion herein, it is helpful to roughly divide the protein into three regions as follows: 1) an amino-terminal domain containing a putative bipartite Nuclear Localization Signal (NLS) sequence [Brar, 2004; Ito, 2004; McBride, 2004; Shinkura, 2004]; 2) a core domain containing a Cytidine Deaminase Catalytic Active Site sequence [Cascalho, 2004; Conticello,

2005; Muramatsu, 1999]; and 3) a carboxy-terminal domain containing a Nuclear Export Signal (NES) sequence (Figure 4) [Brar, 2004; Ito, 2004; McBride, 2004]. Each of these domains has been found to play crucial roles in the regulation and function of the AID molecule [Durandy, 2006; Ta, 2003].

The cytidine deaminase catalytic active site of AID, which is of foremost importance to its functional role in both SHM and CSR, quite tellingly maintains conservation of motifs determined to be important for the deamination function of cytidine deaminases (e.g. the zinc coordination consensus [H/C]XE and PCXXC) [Conticello, 2005; Honjo, 2004; Pham, 2005]. Conticello et al. [Conticello, 2005] also identified another motif (CYX[VI]TW[YF]XS[WS]S) in the active site that appears to be restricted to the AID/APOBEC family. Furthermore, analyses of mutations within the cytidine deaminase catalytic active site of AID support the importance of the conserved cysteines and certain other conserved amino acids for the cytidine deaminase function of the enzyme [Barreto, 2003; Cascalho, 2004; Durandy, 2006; Durandy, 2005; Ramiro, 2003].

Structural modeling of AID continues to rely on predictions from the structures of other deaminases and mutants [Budzko, 2017; King, 2017]. King and Larijani [King, 2017] have recently proposed that the catalytic pocket of AID is inherently unstable and that this instability results in an open state, wherein the catalytic pocket is accessible to substrate binding, and a closed state, wherein the catalytic pocket is “occluded” and therefore inaccessible to substrate. Moreover, they propose that the catalytic rate of AID may be dynamically regulated by controlling access through catalytic pocket closure [King, 2017]. In addition, the differential ability of AID to “anchor” to different substrates (i.e. C versus 5-mC) may also influence catalytic activity and substrate discrimination [Budzko, 2017].

The amino-terminal domain of AID, which contains an alpha helix [Bransteitter, 2004; Cascalho, 2004; Ta, 2003], has been found to play an important role in SHM [Durandy, 2006; Durandy, 2005; Shinkura, 2004]. AID has a net positive charge of +11 at pH 7.0, which is mainly due to a basic amino-terminal domain that is relatively rich in arginine and lysine, a feature that is thought to contribute to the apparent processivity of the molecule [Bransteitter,

2004; Pham, 2005; Pham, 2003], possibly through interaction with putative co-factors (e.g. postulated recruitment proteins) and/or enhancement of affinity for substrate DNA [Bransteitter, 2004; Bransteitter, 2006; Brar, 2004; Cascalho, 2004; Durandy, 2006; Pham, 2003; Ta, 2003].

Mutations in the amino-terminal domain that reduce the charge have been found to variably impact the targeting hierarchy of hotspots within the same sequence context, increase the proportion of mutant clones, and decrease the number of mutations per clone *in vitro* [Bransteitter, 2004; Pham, 2003] and alter or inactivate SHM and CSR *in vivo* [Bransteitter, 2004; Pham, 2005; Shinkura, 2004; Ta, 2003], although these effects cannot be unequivocally attributed solely to charge alterations as the mutations likely impact the protein in other ways (e.g. conformation) [Durandy, 2006; Honjo, 2004]. It is possible that the alterations to SHM *in vivo* may be due in part to disturbances in subcellular localization of the protein or to disruption of protein-protein interactions as a result of the mutations in the putative NLS [Barreto, 2003; Durandy, 2006; Honjo, 2004; Ito, 2004; McBride, 2004; Shinkura, 2004], which has been determined to constitute part of a conformational NLS [Barreto, 2011; Ganesh, 2011; Patenaude, 2009]. That CSR, including point mutation at the switch region, is often not strongly impacted by amino-terminal AID mutations indicates that reduction in SHM is not likely due to concomitant changes in catalytic function or shuttling [Durandy, 2006; Durandy, 2005; Shinkura, 2004]. Interestingly, while the NLS sequence in the amino-terminal domain may have some function in the nucleus and likely plays a role in directing AID into the nucleus, it may not act in an entirely conventional manner to drive the protein into the nucleus [Aoufouchi, 2008; Barreto, 2005; Brar, 2004; Durandy, 2005; Ito, 2004; McBride, 2004].

It has been suggested that AID molecules may act as monomers and/or higher-level multimers (e.g. dimers or tetramers) [Bransteitter, 2004; Chaudhuri, 2004; Dickerson, 2003; Honjo, 2004; Kinoshita, 2006; Li, 2003; Pham, 2005; Ta, 2003], which might provide another layer of regulation, but the exact form involved in SHM and CSR under physiological conditions has not been conclusively established [Durandy, 2006; Durandy, 2005]. Although monomeric

AID is small enough to passively diffuse from the cytoplasm into the nucleus through nuclear pore complexes, which permit passive passage of molecules less than approximately 60 kDa, AID is actively transported out of the nucleus by the ubiquitous CRM1 export pathway that is dependent on the presence of an intact NES in its carboxy-terminus [Brar, 2004; Cascalho, 2004; Ito, 2004; McBride, 2004]. Studies with GFP-fusion proteins have revealed that AID appears to be actively compartmentalized in the cytoplasm and that the NES is instrumental to the rapid shuttling of AID from the nucleus to the cytoplasm [Aoufouchi, 2008; Barreto, 2005; Bransteitter, 2006; Ito, 2004; McBride, 2004; Rada, 2002]. This export is so efficient that AID accumulates in the cytoplasm even when it is fused to an NLS that is known to be conventionally functional [McBride, 2004]. This highlights another early conundrum of AID research: although its DNA target resides in the nucleus, AID is often only observed in the cytoplasm [Aoufouchi, 2008; Ito, 2004; McBride, 2004; Rada, 2002]. But the apparent absence of AID in the nucleus in GFP fusion experiments is likely an artifact due to the sensitivity limits of the detection methods employed as, in the more sensitive chromatin immunoprecipitation assays, AID has been found to be associated with regions undergoing CSR [McBride, 2006; Nambu, 2003; Pham, 2005].

Alteration of the NES in AID leads to disruption of intracellular trafficking that changes its cellular localization from cytoplasmic to nuclear [Aoufouchi, 2008; Ito, 2004; McBride, 2004]. For catalytically active molecules ectopically expressed in fibroblasts, this increase of NES-defective AID in the nucleus is accompanied by an increase in mutation in a target GFP reporter gene (i.e. more mutated clones, more mutations per clone, and more deletions) [McBride, 2004] and in various Ig and non-Ig genes [Aoufouchi, 2008] but SHM in the endogenous IgH gene in B-cells does not appear to be altered [Barreto, 2003; Bransteitter, 2006; McBride, 2004; Ta, 2003] and hotspot preference does not appear to be affected [McBride, 2004].

Although the increased nuclear concentration of NES-defective AID appears to be primarily responsible for the observed increase in mutations,

overexpression of AID with GFP fused to the carboxy-terminus in Ramos cells likewise led to a relative increase in SHM and the proportion of deletions, despite only being detectable in the cytoplasm [Rada, 2002], suggesting that perturbation of the carboxy-terminus (e.g. deletion of the NES or addition of a fusion tag) may disrupt more than shuttling interactions. Further investigation revealed that some alterations of the NES increase the catalytic activity of the enzyme [Barreto, 2003; McBride, 2004], which complicates efforts to assess the contributions of potential co-factors. That this increase in catalytic activity may be due to conformational differences related to NES alteration [Li, 2003] is supported by observations of increased activity when ectopically expressed in E. coli, where AID is not sequestered away from the DNA as it is in eukaryotes, and on ssDNA substrate *in vitro* [Bransteitter, 2004].

Alteration of the NES has profound consequences for antibody affinity maturation [Durandy, 2006]. Mutational studies and examinations of individuals with aberrant immune functions due to antibody affinity maturation deficiencies that preclude or diminish CSR (i.e. heterogeneous HIGMs) revealed a subset, the aforementioned autosomal HIGM2, attributed to various mutations and deletions of the AID gene [Durandy, 2005; Honjo, 2004; Lee, 2001; Li, 2003; Minegishi, 2000; Pham, 2005; Revy, 2000; Stavnezer, 2002; Ta, 2003]. Despite the impairment to CSR, it was observed that some of these mutations do not disrupt AID's catalytic ability and appear to have little or no functional impact on SHM [Bransteitter, 2006; Cascalho, 2004; Durandy, 2006; Durandy, 2005; Ito, 2004; Li, 2003; Pham, 2005; Shinkura, 2004; Ta, 2003]. Further analyses revealed that the mutations that differentially impacted CSR capacity were typically located in the carboxy-terminus of the AID molecule [Ito, 2004; Li, 2003; McBride, 2004; Shinkura, 2004; Ta, 2003] where the NES, which is notably rich in leucine and hydrophobic amino acids [Durandy, 2006; Durandy, 2005; Jacobs, 2001; McBride, 2004], is located. This supports the idea that, in addition to export from the nucleus, the region may be involved in interactions with CSR-specific co-factors [Barreto, 2003; Durandy, 2006; Jacobs, 2001; Li, 2003; McBride, 2004; Pham, 2005; Shinkura, 2004; Ta, 2003].

It appears that AID dimers are still able to form in carboxy-terminal deletion mutants that are not able to class switch [Ta, 2003] and, as noted above, an *in vitro* experiment confirmed that the deaminase function of the molecule was not adversely impacted by truncation at the carboxy-terminus [Bransteitter, 2004; Durandy, 2006; Durandy, 2005; Li, 2003]. Indeed, Barreto et al. [Barreto, 2003] found that, while deletion of the ten carboxy-terminal amino acids (i.e. NES) of AID essentially eliminated CSR in mouse B-cells, mutations in UNG-deficient E. coli were actually increased (~ 10-fold relative to wild-type AID). This has been observed in other experiments [Ito, 2004; Li, 2003] and a similar, although less intense, pattern of SHM increase in fibroblasts and UNG-deficient E. coli was also noted when only one amino acid (F198A) was altered in the carboxy-terminus of AID [McBride, 2004].

Intriguingly, despite the blatant impairment to CSR, not only does carboxy-terminal altered AID (i.e. deletion or mutation) retain SHM capability, it can also still generate mutations in IgH gene switch regions [Barreto, 2003; Durandy, 2006; Li, 2003; McBride, 2004; Pham, 2005]. This indicates that AID catalytic function alone is insufficient to achieve meaningful CSR and provides further support for the hypothesis that, in addition to deamination, AID might also play a regulatory role in directing (e.g. recruiting or positioning) the actions of other molecules involved in CSR (e.g. repair proteins), possibly by participating in a scaffold structure to support a specific resolution (i.e. CSR) of the introduced lesion [Barreto, 2003; Brar, 2004; Durandy, 2006; Li, 2003; Pham, 2005; Poltoratsky, 2004; Rada, 2002; Ta, 2003].

Differential requirement for different domains of AID to bring about SHM and CSR by recruitment of different co-factors handily explains how these different outcomes diverge from a common first step (i.e. C deamination) [Brar, 2004; Cascalho, 2004; Ta, 2003] and differences in data from *in vitro* and *in vivo* cellular and *in vitro* cell-free systems support the use of accessory factors [Barreto, 2003; Durandy, 2006]. Because ectopic expression of AID in non-B-cells of various lineages and in cells of fungi and bacteria leads to hypermutation via C deamination, it has been proposed that the co-factors of AID are

ubiquitously expressed [Durandy, 2006]; however, mutation that occurs as a result of ectopic expression of AID in mammalian non-B-cells does differ from physiologic SHM (discussed above), indicating that perhaps not all co-factors are ubiquitously expressed or equivalently post-transcriptionally modified.

While early studies indicated that the amino- and carboxy-terminal domains of the AID molecule likely help to regulate its function in at least two ways — by regulating subcellular localization through nucleocytoplasmic shuttling and by interacting with various co-factors to differentially promote SHM and CSR [Barreto, 2003; Brar, 2004; Ito, 2004; McBride, 2004; Shinkura, 2004] — later studies revealed that the AID protein is regulated in other ways as well. For instance, it has been observed that the stability of AID-EGFP fusion proteins varies according to their cellular location [Aoufouchi, 2008]. It appears that AID proteins located in the nucleus are preferentially targeted, through ubiquitination, for proteasomal degradation more rapidly than AID proteins located in the cytoplasm [Aoufouchi, 2008]. Indeed, the significantly reduced half-life of NES-defective AID-EGFP (~ 1/3 that of wild-type AID) appears to be related to its inability to exit the nucleus, which presumably makes it more accessible to polyubiquitination. This suggests that differential proteasomal degradation may help to regulate the amount of AID that is available to be actively engaged in AID-mediated processes in the nucleus where its target DNA is located [Aoufouchi, 2008] — a factor which further complicates attempts to ascertain threshold levels of AID required for functional outcomes. Intriguingly, while this pattern of enhanced degradation of nuclear AID occurs in various cell types when AID is ectopically expressed, it appears to occur more rapidly in B-cells [Aoufouchi, 2008], a possible indication of a more deliberate approach to AID regulation within the cells that physiologically express it.

It appears that other post-translational modifications play a role in regulating AID function [Chaudhuri, 2004; McBride, 2006]. In activated B-cells, protein kinase A (PKA) phosphorylation of AID at serine-38, and possibly other sites, appears to enhance participation of AID molecules in SHM and CSR [Bransteitter, 2006; Chaudhuri, 2004; McBride, 2006]. Although they do not

appear to alter the enzyme's deaminase ability, mutations of AID that disrupt phosphorylation at position 38 significantly reduce (>60%) SHM and CSR even when AID is ectopically overexpressed at a level well above physiological (~ 10-fold) [McBride, 2006]. Comparable levels of phosphorylation were also noted when AID was ectopically expressed in non-B-cells and disrupting phosphorylation through mutation at position 38 had the same impact on hypermutation that it did in B-cells [McBride, 2006]. This broadens and bolsters the idea that AID likely uses ubiquitously expressed factors.

That the serine-38-phosphorylated form of AID was found to be preferentially associated with chromatin, despite comprising only a small percentage (<15%) of the total AID in activated B-cells, suggests that phosphorylation may facilitate SHM and CSR by promoting the recruitment of AID to the target DNA [McBride, 2006]. It appears that phosphorylation may help to excite AID activity, at least in part, by promoting interactions with co-factors and/or substrate DNA [Chaudhuri, 2004; McBride, 2006].

One such co-factor may be the ubiquitously expressed replication protein A (RPA), an ssDNA-binding protein [Honjo, 2004] that is essential in DNA replication, recombination, and repair [Barreto, 2005; Binz, 2004; Bransteitter, 2006; Chaudhuri, 2004; Durandy, 2006; Ramiro, 2004]. Recruitment of AID to the Ig genes for SHM and CSR appears to be facilitated by RPA, which has been postulated to form a DNA-binding complex with phosphorylated AID [Chaudhuri, 2004; Ramiro, 2004; Zarrin, 2004]. Binding of the AID-RPA complex to substrate DNA, possibly influenced by RGYW content, appears to lead to deamination of the substrate and then to the dissociation of AID [Chaudhuri, 2004]. The abandoned RPA may then be available to influence activities downstream from deamination such as recruitment of repair molecules to the site [Chaudhuri, 2004; Rada, 2004; Ramiro, 2004] as, in addition to binding to ssDNA and AID, RPA is known to participate in a variety of protein-protein interactions (e.g. UNG2 [Barreto, 2005; Binz, 2004; Chaudhuri, 2004; Durandy, 2006; Durandy, 2005; Honjo, 2004; Kavli, 2005; Kavli, 2007; Rada, 2002; Ramiro, 2004] and RNA pol II

[Chaudhuri, 2004; Nambu, 2003]) that can be differentially regulated by its phosphorylation state [Binz, 2004].

Moreover, Michael et al. [Michael, 2003] found that the introduction of two binding sites for E47 — an important E box-binding transcription factor of B-cells for which binding motifs can be found in Ig, BCL-6, and CD79α and β genes — into the VJ region of an Ig κ transgene increased SHM *in vivo* (i.e. increase in number of mutated clones and mutations per clone) without increasing transcription. This hints that other trans-acting cellular factors likely contribute to recruiting AID to the Ig gene [Michael, 2003]. Attempts to identify direct or indirect interactions with AID or AID-containing complexes are ongoing [Yang, 2006] and an increasing variety of transcription factors (e.g. E2A, PAX5, IRF4, and ETS1) have been found to be able to associate with AID and facilitate recruitment of AID to DNA substrates containing their cognate binding sequences *in vitro* [Grundstrom, 2018]. And it appears that some transcription factors can associate with AID at the IgH gene locus in activated mouse B-cells [Grundstrom, 2018]. But their precise contributions to AID targeting *in vivo* have not been determined.

In addition to DNA-binding transcription factors, AID recruitment to and/or activity at the Ig genes appears to be facilitated by other trans-acting cellular factors, such as transcription elongation factor Spt5, phospho-binding adaptor protein 14-3-3, and spliceosome-associated protein CTNNBL1 [Barreto, 2011; Ganesh, 2011; Yeap, 2019]. Ganesh et al. [Ganesh, 2011] found that the CTNNBL1 appears to interact with the conformational NLS of AID and contributes to its accumulation in the nucleus.

In addition to interactions with trans-acting cellular factors, the subcellular localization of AID and its access to Ig genes appears to be regulated by the cell cycle. Antibody diversification events mediated by AID (i.e. SHM and CSR) appear to be restricted to the G1 phase of the cell cycle [Wang, 2017]. In addition to active import and export, it appears that the dissolution and re-formation of the nuclear envelope during cell division acts as a cell cycle-dependent mechanism to regulate nuclear access of AID [Barreto, 2011; Wang, 2017]. Dissolution of the

nuclear envelope allows AID to redistribute to occupy the same cellular space as the DNA [Wang, 2017]. Re-formation of the nuclear envelope does not appear to result in substantial export and exclusion of AID from the nucleus until early G1, when it is redistributed to the cytoplasm [Wang, 2017]. During the time that AID is colocalized with the DNA, it appears that the conformation of the DNA renders it inaccessible to AID until early G1 — when transcription reactivates — and thereby also contributes to the restriction of SHM and CSR to early G1 [Wang, 2017]. Prolonging the presence of AID in the nucleus beyond early G1 led to increased CSR and mutation in the Sμ and non-Ig gene targets (i.e. off-target mutation) [Wang, 2017].

As AID-mediated deamination has the potential to be both eminently beneficial and disastrously detrimental to the survival of mammals, it is not surprising that AID is tightly regulated [McBride, 2006]. So far, several intricate layers of regulation, including cell-specific, temporally-restricted gene expression, sequestration from the substrate, modulation of function through post-translational modification and co-factor interaction, specific recruitment to target (i.e. Ig genes), and differential protein stability, have been identified. In fact, regulation not only contributes to the direction and restraint of the highly mutagenic AID enzyme but also to its ultimate impact by coordinating downstream processes that recognize and resolve the AID-mediated lesion to promote SHM and CSR within specific regions of the Ig genes.

1.2.7d AID Has Been Implicated In Other Functions

At the inception of this thesis research, AID was known to play an essential role in antibody diversification of the primary and/or secondary repertoire through SHM, CSR, and gene conversion [Arakawa, 2002; Muramatsu, 2000; Muramatsu, 1999]. While the dogma was that secondary repertoire diversification occurred in germinal centers, it has since been established that AID exhibits an expanded expression profile in antibody diversification [Garcia-Carmona, 2018; Patel, 2018]. That is, it appears that AID expression contributes to CSR in B-cells

located outside of germinal centers, an event which likely plays a role in the immune response to bacterial antigens encountered in mucosal tissues [Patel, 2018]. This T-cell-independent response appears to be activated by engagement of the receptors TACI, BCMA, and/or BAFFR with their respective ligands, BAFF or APRIL [Garcia-Carmona, 2018; Patel, 2018]. Mouse and human studies indicate that this intricately complex system of molecules, which plays roles in several aspects of B-cell development, activation, and homeostasis, may have overlapping but distinct functions in different organisms [Garcia-Carmona, 2018].

Additionally, recent research has indicated that AID also contributes to other processes that appear to rely on its activity as a cytosine deaminase on ssDNA. These include epigenetic reprogramming in primordial germ cells and germinal center B-cells and the enforcement of central tolerance during B-cell development [Bochtler, 2017; Kuraoka, 2018; Moris, 2014; Villota-Herdoiza, 2013]. Furthermore, AID has been proposed to play a role in innate antiviral responses, but the mechanism for this activity is unclear [Moris, 2014]. These events are more fully explored in the Discussion.

1.3 The Teleost Humoral Immune System: Antibody Affinity Maturation Capabilities Of A Lower Vertebrate Model Organism — The Channel Catfish

The essential elements of the cognate adaptive immune system, described above in the mammalian paradigm of adaptive immunity, appeared with the advent of the cartilaginous fishes (sharks and rays) [Tort Bardolet, 2003], approximately 500 million years ago [Kumar, 1998], albeit in a more primitive state than that found in mammals [Diaz, 2001; Hsu, 2006; Manser, 2004; Paul, 1999]. Indeed, the lower vertebrates have some obvious differences from the archetypal paradigm of antibody affinity maturation derived from the study of mouse and human [Cain, 2002; Hsu, 2006]. Notably, antibody affinity maturation in lower vertebrates has been found to proceed far more slowly and to achieve only comparatively minor increases in antibody affinity [Cain, 2002; Hsu, 1998]. However, lower vertebrates, except for the placoderms, have flourished for

hundreds of millions of years, which substantiates that their immune systems function adequately for their survival needs [Clem, 1996; Hurley, 2007; Kumar, 1998; Santini, 2009; Tort Bardolet, 2003; Volff, 2005]. In fact, the teleosts alone comprise approximately half of all current vertebrate species [Clem, 1996; Hurley, 2007; Santini, 2009; Volff, 2005]. Yet, in higher vertebrates, similarly low antibody affinity maturation, due to immune defects and/or old-age, is associated with increased morbidity and mortality [Durandy, 2006].

In order to determine the underlying basis for the observed differences in antibody affinity maturation between mammals and lower vertebrates, we explored aspects of this process in a vertebrate model organism, the teleost channel catfish [Clem, 1990; Clem, 1996]. The most recent common ancestor of mammals and ray-finned fishes (Actinopterygii), which include teleosts, existed approximately 450 million years ago [Kumar, 1998; Santini, 2009]. Note that, although teleosts are extant representatives of a class of organisms that arose hundreds of millions of years ago, they have continued to evolve during the interim and, thus, contemporary teleost species are not static snapshots of that ancient point in time [Santini, 2009; Volff, 2005]. Nevertheless, comparative exploration of the differences in the immune responses of mammals and teleosts provides valuable information on the essential requirements of antibody affinity maturation in these organisms.

While zebrafish tend to be the teleost research model of choice, the much larger and longer-lived (~15 years) channel catfish is a particularly good model organism for the study of antibody affinity maturation processes in teleosts for several reasons. As a major freshwater aquaculture focus in the United States, channel catfish is an “economically important” species [Clem, 1984]. Consequently, a good deal is already known about their genetics and physiology including their immune system [Clem, 1996; Clem, 1984; van Ginkel, 1994]. Furthermore, fundamental and applied research is ongoing [Clem, 1996; Clem, 1984; Volff, 2005]. Both B-cells and T-cells — accompanied by antigen presentation, MHC restriction, and allogeneic responses — have been identified in channel catfish [Miller, 1994] and, like other fish [Cain, 2002], they can be

vaccinated to stimulate a T-cell-dependent humoral immune response [Clem, 1996; Iwama, 1996; Vallejo, 1991]. But, perhaps most importantly, while *in vitro* culture of immune cells from other fish species is mostly limited to primary cell culture, channel catfish B-cells, T-cells, and other immune cells (e.g. monocytes) can, for as yet unknown reasons, also be grown in culture as immortal cell lines [Clem, 1996]. Clones can be derived that can be more fully characterized and, typically, manipulated by conventional research techniques (e.g. genetic transformation and mitogen stimulation) [Clem, 1996; Miller, 1994; Miller, 1994; Vallejo, 1991]. However, immortal B-cell lines are sometimes refractory to treatment [Sizemore, 1984], especially to some commercially available cytokines (i.e. mammalian cytokines) [Clem, 1996], probably because some cytokines are not generally well conserved between mammals and teleosts [Tort Bardolet, 2003].

1.3.1 Catfish Primary Antibody Repertoire Generation Approximates Mammalian

As a bony fish, channel catfish is a representative of some of the earliest organisms to have a translocon IgH gene organization similar to that seen in mammals [Clem, 1996; Hsu, 1998; Hsu, 2006; Magor, 1999; Magor, 1994] and their B-cell development and maturation appears to be likewise similar to that described above in the mammalian paradigm (Figure 6) [Miller, 1994; Zwollo, 2005], but there are some relevant differences that may contribute to their apparently reduced capacity for antibody affinity maturation [Tort Bardolet, 2003]. For example, teleosts lack bone marrow, which is the main site of hematopoiesis in mammals; instead, hematopoiesis occurs in the kidney, which is a site of manifold functions, including excretory (i.e. renal), endocrine, and lymphohematopoietic functions [Fange, 1985; Iwama, 1996; Tort Bardolet, 2003; Zwollo, 2005]. The anterior portion of the kidney is mainly hematopoietic tissue with limited renal function [Fange, 1985; Tort Bardolet, 2003; Zwollo, 2005]. It also appears to harbour long-lived plasma cells [Zwollo, 2005]. The posterior portion of the kidney has greater renal function but also acts as secondary

lymphoid tissue that serves as a site for activated B-cells undergoing proliferation and antibody affinity maturation [Tort Bardolet, 2003; Zwollo, 2005].

Generation of channel catfish primary antibody repertoire during B-cell development approximates that previously described for mammals, with RAG-mediated recombination of the diverse V, D, and J sub-exons, junctional diversity, allelic exclusion, expression of both IgM [Miller, 1994; Ohta, 2006] and IgD [Ohta, 2006; Wilson, 1997] isotypes through differential splicing of IgH mRNA transcripts, and pairing of IgL and IgH polypeptides to create vast potential for diversity of binding sites for antigen recognition in a polyclonal B-cell population [Hsu, 1998; Hsu, 2006]. And, also like mammals, channel catfish and other teleosts exploit avidity strategies to enhance the affinity of their primary antibody repertoire by producing tetrameric serum IgM [Cain, 2002; Clem, 1996; Hsu, 2006; Iwama, 1996; Kaattari, 2002; Tort Bardolet, 2003].

Interestingly, although there are no additional IgH gene constant regions (e.g. the γ, ε, and α found in mouse and human) downstream of Cμ and Cδ [Hsu, 1998; Magor, 1999], many teleost fish species have been found to possess a third IgH isotype that appears to be specialized for mucosal tissues [Hsu, 2006; Patel, 2018; Sunyer, 2013; Wakae, 2006]. This third IgH isotype, called IgT (or IgZ), is expressed as a result of alternative VDJ rearrangement during B-cell development rather than through CSR after B-cell activation [Patel, 2018]. Analysis of channel catfish indicates that it does not possess a third IgH isotype [Edholm, 2010].

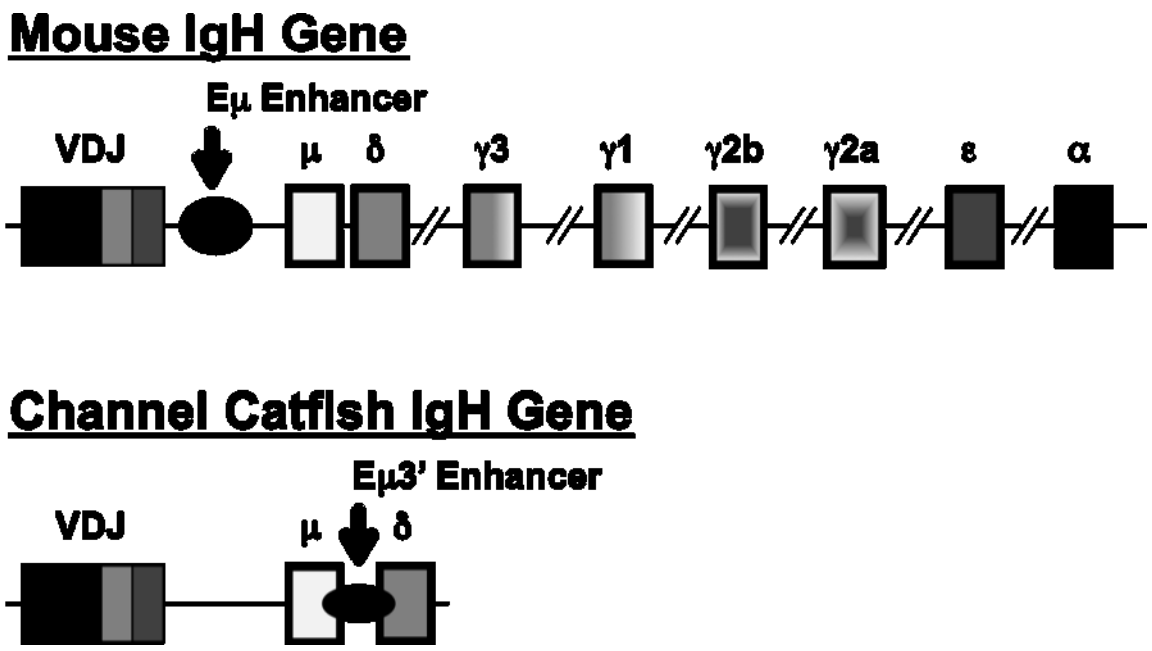


Figure 6. Mouse versus channel catfish immunoglobulin heavy chain genes.

As in the mouse, RAG-mediated rearrangement of V, D, and J sub-exons gives rise to the complete VDJ exon in the channel catfish IgH gene. But the channel catfish IgH gene differs from that of the mouse in two important ways. First, the channel catfish IgH gene lacks additional constant regions downstream of the germline μ and δ constant regions, which precludes the occurrence of CSR. Second, the channel catfish IgH gene transcriptional enhancer, Eμ3', is located amidst the constant regions, a position that is incompatible with CSR as it would be excised. Diagrams are not drawn to scale and constant region exons have been depicted as single boxes for clarity. Adapted from [Magor, 1999].

1.3.2 Catfish Exhibit Meager Antibody Affinity Maturation

In lower vertebrates, repeated exposure to antigen results in only a slight increase in antibody affinity and memory is generally considered to be weak, as successive responses don't typically much exceed the primary response in various assessment metrics (e.g. kinetics, antibody affinity, or titer) [Cain, 2002; Hsu, 1998]. Studies in teleost fish [Cain, 2002; Kaattari, 2002] indicate that they are able to enhance the relative affinity of their antibody response to haptens by approximately 2 to 3-fold [Cain, 2002]. But this enhancement of antibody affinity, which requires months to develop even with repeated exposure to antigen [Cain, 2002; Kaattari, 2002], seems minuscule when compared to that of mammals which, while generally considered to increase in the range of greater than 10-fold [Cain, 2002] to about 100-fold [Foote, 1995], has been measured at as much as 30,000-fold [Wedemayer, 1997] for certain monoclonal antibodies.

It is unclear whether this limited affinity enhancement arises from the preferential proliferation of pre-existing higher-affinity antibody-expressing B-cell clones (i.e. selective enrichment and clonal expansion) or from relatively inefficient conventional antibody affinity maturation processes (e.g. reduced Ig SHM and/or less stringent selection) [Cain, 2002; Kaattari, 2002]. Regardless of the strategy employed, because channel catfish serum IgM is a tetramer (∴ eight antigen-binding sites) [Clem, 1996], even a slight increase in affinity may result in a far more appreciable increase in avidity [Cain, 2002; Kaattari, 2002; Torres, 2008].

Although the actual physiological impact of antibody affinity differences can be difficult to ascertain across species, for comparative studies, affinity changes provide a necessary standard measure by which to compare the effectiveness of adaptive immune responses. But there are some important caveats to the use of affinity increases as the sole indicator of the effectiveness of immune responses. One caveat is that the number of B-cells able to contribute to the response to a specific antigen, the rate and magnitude of their amplification, and/or the amount of antibody each cell produces may compose a different proportion of the total

population in different organisms. For example, the IgM secretion rate appears to differ between channel catfish and mammals [Miller, 1994]. The rate of IgM secretion measured from the channel catfish immortal B-cell line 1B10 is approximately 7000 molecules/minute/cell whereas estimates from mammals (mouse and human) range from 2000 to 7000 molecules/second/cell, an approximately 17- to 60-fold difference [Miller, 1994]. Moreover, several antigen metrics (e.g. type of antigen, dose, context, and exposure route and duration) can influence both the recruitment of B-cells into the response and the extent of the response [Batista, 2000; Paus, 2006; Tarlinton, 1998]. Another caveat is that measurements of absolute affinity or relative affinity increases of serum antibody don't reflect the *in vivo* antigen-binding context in different organisms where other elements (e.g. innate immune mechanisms) may be differentially enhanced to compensate for perceived shortcomings in antibody affinity maturation [Tort Bardolet, 2003]. Indeed, some compensatory elements may actually contribute to limiting antibody affinity maturation by participating in cooperative co-binding of antigen, along with BCR, at the surface of B-cells, thereby altering the B-cells' ability to thoroughly discriminate antigen based on BCR affinity (e.g. limit diffusion of antigen and/or alter signaling thresholds). An example of this effect of coligation was briefly described in the mammalian section where it was noted that studies indicate that, relative to its absence (i.e. a defect), the presence of CR2 exerts a damping influence on the potential antibody affinity of the resultant population of plasma cells [Tarlinton, 2000]. Because BCR engagement of antigen is important for the functional selection of B-cells during antibody affinity maturation processes, B-cells are not likely able to discriminate between affinities that exceed some maximal level of intracellular signaling, which may be achieved at lower affinities in different organisms or through cooperative binding [Batista, 2000; Tarlinton, 2000]. Although competitive binding (i.e. discrimination through acquisition) may help to enhance antibody affinity slightly beyond the affinity ceiling that would otherwise be imposed by reaching some maximal level of intracellular signaling [Batista, 2000], binding of new high-affinity serum antibody to immune complexes [Tarlinton, 2000] may also help to enforce the affinity

ceiling and modulate the response. The hypothesis that maximal intracellular signaling is achieved before, and so suppresses the development of, maximal antibody affinity is supported by the observation that, even in mammals, the *in vivo* “affinity ceiling” is well below the potential binding-affinity maxima for most antibodies [Batista, 2000; Foote, 1995; Kaattari, 2002]. This suggests that there is a point at which no more meaningful increases can be achieved, at least within the context of the system [Batista, 2000; Foote, 1995; Kaattari, 2002]. But it is also possible that the affinity ceiling is due to factors other than to certain maximal intracellular signaling as there is some indication that too strong a signal may cause centrocytes to undergo apoptosis [Tarlinton, 2000].

Interestingly, in mammals and the amphibian *Xenopus*, there is a difference between IgM and class switched Ig such that IgM does not achieve as high a level of affinity maturation as the others (i.e. IgM has a lower affinity ceiling) [Cain, 2002; Hsu, 1998; Kaattari, 2002]. That IgG has a higher frequency of mutations than IgM suggests that affinity differences between antibody isotypes may be partly due to differences in cis-acting genetic elements associated with the different constant regions as these might influence mutation rates [Green, 1998; Papavasiliou, 2002; Peters, 1996].

Another caveat of using comparative antibody affinity studies as the sole indicator of the effectiveness of adaptive immune responses is measurement accuracy. Due to assumptions (e.g. disregard of on-rates versus off-rates) and limitations of the methodologies used to measure antibody affinity, measurements of the affinity of teleost antibodies may be less accurate than those of mammalian antibodies [Cain, 2002]. Several environmental factors (e.g. ionic strength, temperature, pressure, and pH) can influence the binding affinity of antibodies by their effects on both the antibody and antigen — and not necessarily in equivalent ways [Foote, 1995; Torres, 2008].

The teleost physiological context may not be well replicated in antibody affinity measurement experiments, which often use ‘mammalian’ binding parameters [Burton, 2002; Cain, 2002]. For instance, relative to humans, channel catfish have a lower body temperature, lower isotonicity, and a higher blood pH,

which changes slightly with temperature (pH 7.79 at 25°C and pH 7.92 at 15°C) [Burton, 2002]. But the problems with measurement accuracy have been somewhat mitigated by examining the change in affinity rather than absolute affinity and this has confirmed that the increase in antibody affinity is low in fish even with repeated antigen exposures [Cain, 2002].

If conventional antibody affinity maturation processes do occur in teleosts, then several factors, including the potential intrinsically low affinity ceiling mentioned above, may contribute to the limited antibody affinity maturation observed in teleosts, most likely by disrupting or altering selection and/or SHM. One of these factors might be temperature. As ectothermic vertebrates, the body temperature of teleosts is dependent on the ambient temperature of their environment [Tort Bardolet, 2003]. Because they must operate physiologically at the temperature of their particular environment, their immune systems often have to operate effectively over a wider range of temperatures than in mammals [Tort Bardolet, 2003]. Indeed, many teleost species, including channel catfish, reside in aquatic environments that can undergo seasonal temperature fluctuations of greater than 20°C and where even the highest temperatures are still several degrees below that of mammalian body temperature [Tort Bardolet, 2003]. Body temperature has an impact on several aspects of channel catfish physiology (e.g. membrane fluidity [Bly, 1986; Clem, 1984; Lin, 1992], blood pH [Burton, 2002], metabolism kinetics, cell division time [Hsu, 1998], and protein flexibility, conformation, and binding [Hsu, 1998]) and various studies have demonstrated the negative (i.e. immunosuppressive or non-permissive) impact that colder temperatures (approximately 17°C and below) have on the immune systems of these organisms [Clem, 1984; Hsu, 1998], especially T-cell-dependent responses (i.e. reduced help) [Bly, 1986; Clem, 1996; Lin, 1992; Tort Bardolet, 2003].

Histological differences in secondary lymphoid organs may also contribute to the limited antibody affinity maturation observed in teleosts as lymphoid functions within their tissues do not appear to be as well partitioned as in higher

vertebrates [Fange, 1985; Tort Bardolet, 2003]. Channel catfish do not have lymph nodes, but they do have a spleen, mucosa-associated lymphoid tissue (MALT), and, as noted above, the kidney also operates as a secondary lymphoid organ [Fange, 1985; Iwama, 1996; Tort Bardolet, 2003; Zwollo, 2005]. Activated B-cells aggregate in these organs but, like other lower vertebrates (sharks and amphibians), channel catfish lack histologically obvious germinal centers — structures believed to be essential for efficient antibody affinity maturation in the warm-blooded higher vertebrates [Cain, 2002; Diaz, 1998; Diaz, 2001; Fange, 1985; Hsu, 1998; Kelsoe, 1996]. If they truly lack germinal centers, channel catfish may not be able to carry out efficient selection, a key determinant of the robust antibody affinity maturation observed in mammals.

Despite the absence of identifiable germinal centers, that a small increase in antibody affinity (i.e. 2 to 3-fold) has been measured in trout indicates that some amount of selection for B-cells bearing BCR of higher affinity does occur in teleosts [Cain, 2002; Kaattari, 2002]. Ultimately, even in germinal center environments, selection relies on intrinsic cellular mechanisms of meeting an antigen interaction threshold coupled with appropriate help, which can still occur in the absence of canonical germinal centers despite the diminished efficiency. Indeed, although mice deficient in either TNFR family receptors or their ligands (e.g. lymphotoxin-alpha) typically don't form germinal centers or have FDC they still exhibit some antibody affinity maturation [Diaz, 2001; Hannum, 2000; Matsumoto, 1997]. Despite the absence of obvious germinal centers, nurse sharks appear to exhibit some antigen-dependent selection of their novel new antigen receptor (IgNAR) [Diaz, 1998; Diaz, 2001] and *Xenopus* are able to increase their antibody affinity (less than 10-fold) in an antigen-dependent manner similar to mammals [Cain, 2002; Diaz, 1998; Hsu, 1998]. However, although the antibody affinity increase in *Xenopus* suggests that selection occurs in these organisms, an analysis of the FRs and CDRs revealed that the R/S ratios in both regions approximated that of random mutation, which is deemed to be inconsistent with selection [Diaz, 1998; Hsu, 1998].

Assuming that channel catfish B-cells rely on the same sorts of interactions for selection as mammalian B-cells during antibody affinity maturation, without a defined assembly site in which to concentrate all the requisite participants they may be diluted in the tissues such that the acquisition of antigen and T-cell help, which are central to selection in higher vertebrates, may become less about affinity-based competition among B-cells and more about random encounter in the loose environment, leaving the B-cells unable to achieve a heightened response [Hsu, 1998]. In this scenario, one would necessarily expect there to be an extended window of time for the B-cells which have acquired antigen to also acquire T-cell help, especially given the lower temperature of fish, which raises two important implications: 1) apoptosis of B-cells may not play a strong role here (e.g. T-cell help may be less about B-cell survival per se and more about maintenance of the response, possibly through increased B-cell proliferation), which would likely reduce the potential affinity of the response as it does in mice in which apoptosis of germinal center B-cells has been blocked [Tarlinton, 2000]; and 2) there may be a necessarily slower turnover of MHC:peptide complexes, which could further diminish the potential for enhancement of B-cell proliferation based on affinity discrimination as persistence of antigen on B-cells bearing BCR of lower affinity would negate the ability of B-cells bearing BCR of higher affinity to preferentially engage T-cell help as a direct result of their enhanced ability to acquire antigen (i.e. mimics antigen abundance [Kaattari, 2002]).

Indeed, inefficient T-cell help is thought to be one of the reasons that IgM (i.e. unswitched IgH) remains prevalent in primary and subsequent responses in frogs because the balance can be shifted to class switched IgY by either extending the duration of antigen exposure or increasing specific T-cell activity through artificial means (e.g. by increasing their temperature or through the introduction of ‘primed’ T-cells), surprisingly without enhancing affinity [Hsu, 1998].

It is thought that channel catfish may possess “primitive analogues” of germinal centers in the form of melanomacrophage cells, which can be found, individually or in clusters (i.e. melanomacrophage centers), to co-occur with lymphocytes [Agius, 2003; Fange, 1985; Steinel, 2017]. In higher vertebrates,

antigen-trapping by FDCs is believed to be an important component of antibody affinity maturation and teleost melanomacrophages trap antigen on their surface for a long time in a manner reminiscent of the FDCs found in mammalian germinal centers [Agius, 2003; Fange, 1985; Magor, 2015; Steinel, 2017; Tort Bardolet, 2003]. As phagocytic cells, melanomacrophages also operate similar to mammalian tingible body macrophages, another cell type found in germinal centers, to remove apoptotic cellular debris, an activity that causes them to accumulate various pigments, primarily lipofuscin/ceroid but also melanin and haemosiderin [Agius, 2003; Diaz-Satizabal, 2015; Fange, 1985; Magor, 2015; Steinel, 2017; Tort Bardolet, 2003]. Melanomacrophage cells are easy to visualize microscopically due to their pigment content, which causes them to appear brown under bright field and to autofluoresce under UV and various fluorescence channels (Figure 7) [Magor, 2015; Saunders, 2010]. Teleosts have a variety of pigment-containing cells and a number of features distinguish the melanomacrophage cell as a cell type that is distinct from the more prominent non-phagocytic, melanin-containing melanophore, which is a teleost sub-type of melanocyte [Braasch, 2008; Osawa, 2008]. Melanomacrophage cells, which were determined to be tissue-resident macrophages (i.e. derived from embryonic precursors rather than bone-marrow derived circulating mononcytes) subsequent to the research described herein (see Discussion), reside in a variety of immune and non-immune tissues where they may carry out different functions, such as clearance of senescent red blood cells [Agius, 2003; Diaz-Satizabal, 2015; Fange, 1985; Ginhoux, 2016; Steinel, 2017]. The melanocytes of teleosts do exhibit some differences from those of mammals but their developmental pathway appears to be quite conserved [Braasch, 2008; Lorin, 2018; Mort, 2015; Osawa, 2008]. As in mammals, teleost melanocytes exhibit a highly dendritic morphology and, although their ontogeny is still being investigated, appear to originate primarily from neural crest cells during embryogenesis via self-renewing melanoblasts that migrate to and then reside in the skin [Braasch, 2008; Djurdjevic, 2015; Higdon, 2013; Lin, 2007; Lorin, 2018; Mort, 2015; Osawa, 2008; Taylor, 2011]. Melanocytes produce melanin via a tyrosinase-based biosynthesis

pathway [Braasch, 2008; Lin, 2007; Mort, 2015; Osawa, 2008]. Despite conflicting reports, it appears that melanomacrophage cells may not use a tyrosinase-based pathway for biosynthesis of melanin [Agius, 2003; Diaz-Satizabal, 2015; Iwama, 1996]. While it has been surmised that melanomacrophage cells may acquire their melanin through phagocytosis of other cells [Agius, 2003; Diaz-Satizabal, 2015; Iwama, 1996], in mammals and some other tetrapods (e.g. frogs) there is a precedent for production of a melanin-containing molecule, neuromelanin, by an apparently non-tyrosinase-based biosynthesis pathway as neurons of the substantia nigra appear to produce this melaninic molecule without tyrosinase [Zecca, 2001; Zucca, 2014]. Intriguingly, the neuromelanin in these neurons is reported to be associated with lipofuscin and has been postulated to be a mechanism to protect them against oxidative stress that would otherwise be generated due to their high content of iron and other metals [Zecca, 2001; Zucca, 2014]. Notably, high iron content, in the form of haemosiderin, is a feature which is likewise common to the melanomacrophage cells in which melanin has also been postulated to provide protection from oxidative stress generated by degradative activities associated with phagocytosis [Agius, 2003; Diaz-Satizabal, 2015; Iwama, 1996].

In addition to the antigen-trapping and phagocytic activities of melanomacrophage cells, they may exhibit other characteristics that are reminiscent of germinal center-like cell aggregates. Much like the ectopic germinal centers that occur in mammalian diseases (e.g. in synovial joints during arthritis) [Park, 2005], “adventitious” melanomacrophage cell clusters may also occur at sites of chronic inflammation [Agius, 2003] — although whether these clusters contribute positively or negatively to the response is unknown. Interestingly, in ectopic germinal centers of arthritis patients, synoviocytes apparently exhibit some FDC characteristics that include binding interactions with “germinal center B-cells and prevention of apoptosis” [Park, 2005], which suggests that an FDC role can be played by other cell types, at least in mammals.

The finding that, during T-cell-independent responses in mammals, CSR can occur in B-cells that are located outside of defined germinal centers in the course of a normal physiological immune response [Garcia-Carmona, 2018; Patel, 2018] further supports that germinal centers are not definitively necessary for AID-mediated antibody diversification events. It also hints at the possibility that an FDC-type role may not be entirely necessary for these events. Of course, the challenges presented by SHM (e.g. higher risk of autoimmunity) differ from those of CSR and may, therefore, render it less suitable to occurrence outside of what is perceived to be a relatively more structured environment.

The effects of selection may also be diluted by the expanded functional role of teleost B-cells, which are more numerous in the blood of fish than B-cells in the blood of mammals and likely help to compensate for lower antibody affinity maturation in teleosts [Li, 2006]. Teleost B-cells are highly phagocytic [Li, 2006] and the uptake of foreign material in a potentially antigen-independent manner by antigen-activated B-cells may lead to antigen presentation and survival or proliferation signals disconnected from BCR affinity. Indeed, the form of antigen is important in uptake and affinity discrimination and presentation in mammals [Batista, 2000].

But selection, whether in conventional germinal centers or otherwise, needs suitable material to select and it seems that teleost fish may not be able to provide it, which suggests that inadequacies in trans-acting cellular factors (e.g. AID) and/or cis-acting genetic elements (e.g. IgH gene transcriptional enhancer) may also contribute to the limited antibody affinity maturation observed in teleosts. Importantly, the status of AID (i.e. presence and functional capability of a homologue), discovered three years prior to the inception of this research, was unknown in teleosts and other lower vertebrates. However, replicative DNA polymerases and DNA damage tolerance and repair enzymes (e.g. UNG) arose early in evolution and are known to be fairly well conserved between teleosts and mammals [Diaz, 1998; Honjo, 2004; Rada, 2002], although there may be some species-specific differences in their activity levels and substrate preferences [Kavli, 2005].

Despite considerable similarity to the mammalian IgH gene, there are two important differences in the channel catfish IgH gene (Figure 6) that preclude the occurrence of CSR and so impact the potential for secondary antibody repertoire development through antibody affinity maturation: 1) there are no additional IgH gene constant regions (e.g. the γ, ε, and α found in mouse and human) [Hsu, 1998; Magor, 1999]; and 2) the IgH gene transcriptional enhancer differs significantly in its location, content, and organization [Magor, 1999; Magor, 1994]. The barrier to CSR created by the lack of extra constant regions is obvious: CSR cannot occur when there are no additional constant regions to which to switch. Perhaps less evident is the way in which the differences in the channel catfish IgH gene transcriptional enhancer may also thwart CSR [Magor, 1999].

In mice, the IgH gene Eμ intronic transcriptional enhancer, which is compact in its organization, is located in the intron between the most 3' J sub-exon and the first constant region exon [Magor, 1999]. Its position 5' of the μ switch region means that the transcriptional enhancer does not get excised during CSR events [Magor, 1999]. In the channel catfish, the IgH gene transcriptional enhancer, which is more diffusely organized than its mouse counterpart and recruits different factors, is located 3' of the Cμ transmembrane 1 (TM1) region through the TM2 region and into the intron 5' of the first Cδ constant region exon [Magor, 1999; Magor, 1994]. These differences do not appear to affect its function, as it has been empirically demonstrated to enhance transcription just as well as the mammalian IgH gene transcriptional enhancer, but its location would be excised were CSR to occur [Magor, 1999; Magor, 1994].

Although the absence of CSR in teleosts effectively eliminates what is seen as a key element of antibody affinity maturation in mammals (i.e. HIGM results from CSR defects in mammals), it is difficult to estimate the true significance of the absence of CSR in the channel catfish as it actually need not adversely impact affinity improvement per se in a system that has not evolved to exploit partitioning of effector functions through CSR.

In addition to lacking CSR, teleosts have an “apparent lack of SHM” [Cain, 2002], but their ability to hypermutate is actually unknown [Diaz, 1998; Magor, 1999]. It is difficult to predict precisely how the differences in the location, content, and organization of the channel catfish IgH gene transcriptional enhancer might affect its role in SHM. In mammals, the role of the IgH gene transcriptional enhancer in SHM appears to be manifold: it drives transcription of the gene, which is important for access; it appears to serve as the 3’ boundary of SHM, which helps to constrain mutation to the VDJ region; and it may contain cis-acting genetic elements involved in the recruitment of trans-acting cellular factors that contribute to SHM.

Although the channel catfish IgH gene transcriptional enhancer drives transcription as well as its mammalian counterpart [Magor, 1999; Magor, 1994], which suggests that it should be able to provide the mutational machinery with suitable access to the IgH gene, that it recruits different factors to enhance transcription [Magor, 1994] means that recruitment of the SHM machinery may be affected. Additionally, its location 3’ of the Cμ constant region, which is considerably further from the VDJ exon than in mammals, indicates that it may not focus mutation to the VDJ region as effectively as the mammalian IgH gene Eμ intronic enhancer [Magor, 1999] or shield the Cμ constant region from incidental mutation. The latter raises the prospect that SHM could possibly occur in the Cμ constant region exons, although their distance from the IgH gene promoter renders it unlikely.

Despite the lack of SHM data from teleost fish, there is strong genetic evidence for the occurrence of SHM in sharks (IgM and IgNAR) and both SHM and CSR in amphibians [Diaz, 1998; Diaz, 2001; Hsu, 1998; Hsu, 2006; Kenter, 2004; Magor, 1999; Neuberger, 1995; Ohta, 2006; Papavasiliou, 2002; Zarrin, 2004], although there are species differences (e.g. *Xenopus* switch regions are AT-rich) [Diaz, 2002; Kenter, 2004; Neuberger, 2003; Zarrin, 2004]. That SHM not only occurs in sharks and amphibians but also exhibits a strong GC bias (except in IgNAR) and favours AGY hotspots [Diaz, 2002; Diaz, 1998; Diaz, 2001; Hsu, 1998; Papavasiliou, 2002] suggests that mutation in the Ig genes of

these organisms may be instigated by an AID homologue but resolved slightly differently. Indeed, the major similarities in the SHM observed in organisms from various vertebrate classes suggest that the same primary mechanism is used while the minor differences indicate some elements may be unique or absent [Diaz, 1998]. Furthermore, because sharks and amphibians are from classes that evolutionarily flank the ancestor of bony fish in the vertebrate lineage, the occurrence of SHM in these organisms also suggests that teleosts may possess an AID homologue [Hsu, 1998; Kumar, 1998].

That SHM has been identified in the Ig genes of sharks and amphibians indicates that the absence of a canonical germinal center is not a barrier to the occurrence of processes that modify Ig genes [Diaz, 1998; Diaz, 2001; Hsu, 1998]. From the presence of SHM, it is reasonable to hypothesize that some form of selection mechanism likely exists in these organisms. Firstly, SHM of the Ig genes of activated B-cells is a risky strategy. Even in mammals, which are considered paragons of selection during secondary repertoire generation, SHM can contribute to the development of autoimmunity [Durandy, 2005]. That antibody-dependent autoimmunity can be experimentally induced in a teleost fish (i.e. autoimmune orchitis in rainbow trout) suggests that lower vertebrates are also susceptible to such detrimental outcomes [Iwama, 1996]. Secondly, SHM of the Ig genes would be purposeless without some means to promote improved clones. Indeed, as noted above, there is some evidence for antigen-dependent selection in nurse shark (IgNAR) [Diaz, 1998; Diaz, 2001] and *Xenopus* [Cain, 2002; Diaz, 1998; Hsu, 1998]. But whether the apparent selection in these organisms employs the same mechanisms that operate in mammalian germinal centers (e.g. rescue from apoptosis) or relies on other mechanisms (e.g. preferential amplification) has not yet been determined.

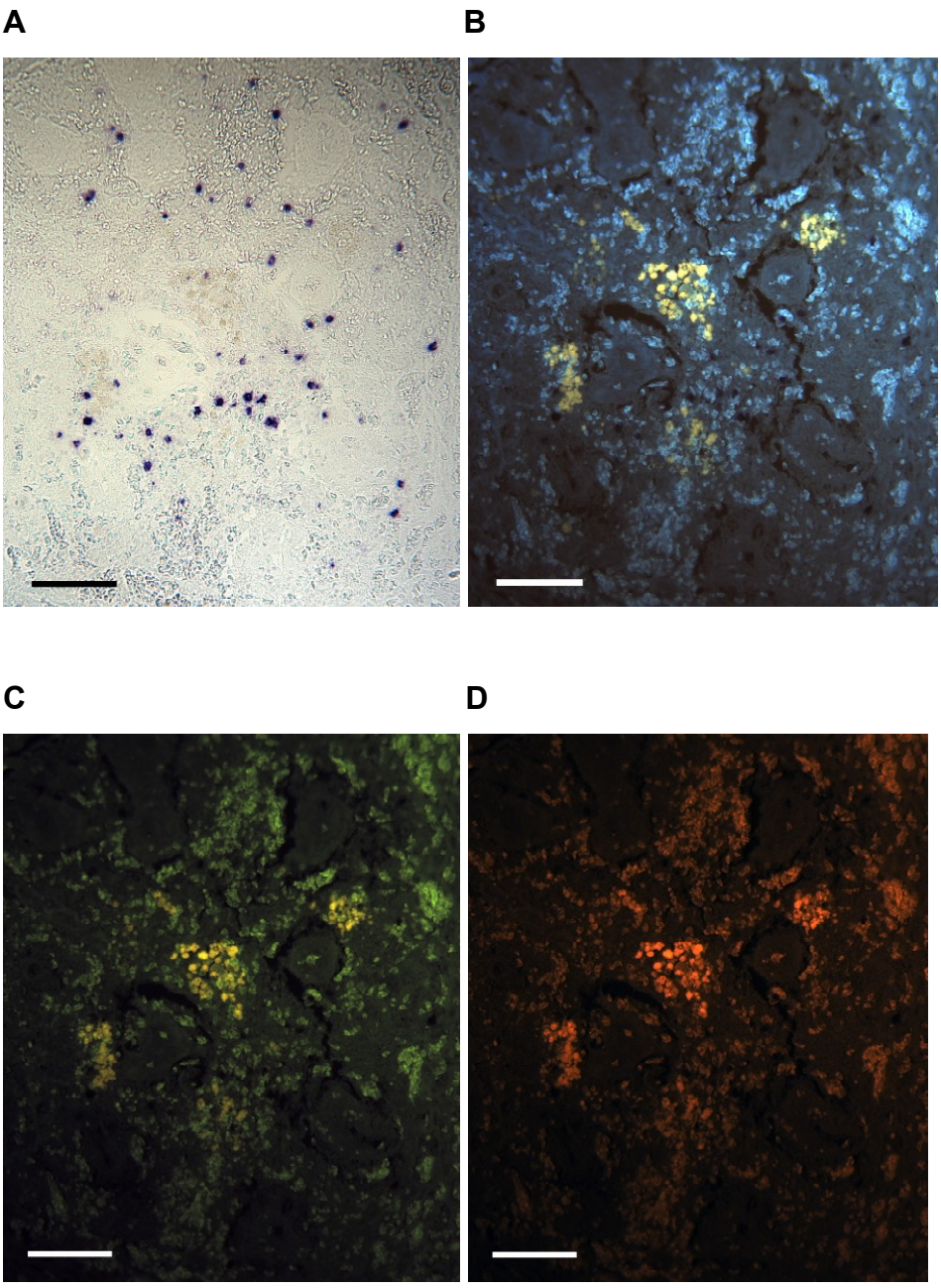


Figure 7. Melanomacrophage cells contain pigments that render them visible during microscopy without additional staining.

In these images of the same region of a 5 µm thick tissue section from channel catfish spleen, the pigment content of melanomacrophage cells causes the cells to appear very pale brown under bright field microscopy (A) and to variably autofluoresce under UV (B) and green (C) and red (D) fluorescence channels. These images are typical but alterations to the image capture parameters and/or

treatment of the tissue sections can impact the relative intensity and hue of the autofluorescence of melanomacrophage cells and surrounding tissue. Note how a comparison between images captured under various light channels can allow for better resolution of stained cells from melanomacrophage cells. In this example from ISH for IgH mRNA, IgH-positive cells stained with blue-purple formazan chromogen (i.e. from NBT/BCIP) are most visible in image A whereas the melanomacrophage cells are most visible in images B-D with the best contrast in images B and C. Scale bars represent 50 μm.

1.3.3 Determination Of Catfish AID Status Would Resolve A Crucial Uncertainty

Because, in higher vertebrates, AID is essential to SHM and is expressed within the confines of the germinal center [Muramatsu, 1999], determination of whether channel catfish express an AID homologue would establish conclusively whether the limited antibody affinity maturation observed in teleost fish may be due in part to a lack of AID-mediated processes (i.e. SHM). Additionally, the identification of a channel catfish AID homologue would provide a means to assess their SHM capacity and serve as a marker to determine if they have a genuine germinal center analogue. These avenues of inquiry can be variously addressed in tissues from channel catfish specimens or their immortal B-cell lines.

In the mammalian paradigm of adaptive immunity presented above, expression of the AID gene occurs as part of the immune response to antigen challenge and it must be expressed above an undefined threshold [Zhang, 2001] before mutations are seen. Some mammalian B-cell lines can be induced to express AID [Muramatsu, 1999] while others express AID constitutively (e.g. Ramos [Zhang, 2001]). As in higher vertebrates, many teleost immune genes are only expressed in response to antigen challenge *in vivo* or to substances (i.e. mitogens) that initiate or mimic similar signaling outcomes *in vitro* [Clem, 1996]. Thus, because AID is expressed at a relatively low level in mammalian B-cells, to ensure the highest possible level of expression of any potential AID homologue in the channel catfish specimens and immortal B-cell lines examined, they must each be treated in a manner known to stimulate their adaptive immune response and upregulate the expression of immune genes.

To maximize T-cell-dependent humoral immune responses in tissues, the channel catfish specimens, which were not held in sterile conditions and all displayed obvious signs of immune activation (e.g. inflammation), were vaccinated repeatedly with bovine serum albumin, a non-glycosylated protein, prior to harvest (detailed in Materials and Methods). And, to upregulate immune

genes, the immortal B-cell lines were treated with various mitogens (detailed in Materials and Methods).

Explorations of AID homologue expression, including the existence of germinal center analogues, can be rather simply addressed in catfish tissues, but the current paucity of genetic data makes the analysis of any *in vivo* SHM rather difficult. The channel catfish immortal B-cell lines, which express a defined IgH sequence, provide an excellent means by which to search for SHM (i.e. an indirect indicator of AID function) in the IgH gene of fish cells in the absence of antigen-dependent selection, which can skew mutational outcomes. While mitogenic activation of these cells might be enough, if an AID homologue is present, to initiate SHM for preliminary analysis, this type of chemical-mediated upregulation of the expression of a possible endogenous AID gene is not ideal for several reasons, including inconsistent levels of gene expression, non-specific upregulation of genes, and it does not allow for comparative examination of AID from other organisms or mutant AID. In contrast to mitogenic stimulation, ectopic expression of AID, through integration of an AID expression construct from which expression can be controlled to prevent global genome mutation during culture (i.e. Tet-Off system), is a reliable means of surpassing the expression threshold in a way that is not only consistent but also specific to any AID gene of interest. In addition to allowing expression of AID from various organisms, such a system would offer the flexibility to examine AID-mediated processes in reporter transgenes (e.g. assess CSR potential).

In summary, the intent of this research is to identify a channel catfish AID homologue and to assess its tissue expression pattern *in vivo* in channel catfish and its functional capacity at the endogenous IgH gene *in vivo* in a channel catfish immortal B-cell line. In addition to gene organization and protein homology, a tissue expression profile consistent with a role in antibody affinity maturation (i.e. secondary lymphoid tissues) and a demonstration of functional activity (i.e. SHM) in the endogenous IgH gene of channel catfish cells following expression of native AID should help to verify that the sequence identified in channel catfish constitutes an AID homologue. That is, it contributes to antibody

affinity maturation processes rather than to other functions in fish (e.g. innate antiviral responses like mammalian APOBEC3s). The capacity of channel catfish immortal B-cells to support mutation of the endogenous IgH gene or reporter constructs when transgenically supplied with AID from various organisms will also be assessed. The addition of appropriate reporter constructs will also allow for assessment of CSR capacity in channel catfish immortal B-cells. This will allow for a comparison with mutation data obtained from complementary studies in which AID homologues from teleosts and other organisms were expressed in the context of mammalian cells. And the AID homologue will be used as a marker to identify B-cells that are, presumably, in a state equivalent to those found in the germinal centers of mammals to establish whether a discernible cellular organization operates in teleosts in lieu of conventional germinal centers during responses to T-cell-dependent antigens. Since mammalian germinal centers are sites of B-cell selection, the identification of such a discrete organization as a site of SHM in channel catfish would plausibly suggest that some form of selection likely accompanies Ig gene modification in teleost fishes. However, it would not be direct proof of selection and additional analysis of such sites would be required in the future to definitively determine whether selection of hypermutated B-cells contributes to teleost antibody affinity maturation.

2. Materials And Methods

2.1 General Maintenance And Preparation Of Research Specimens

2.1.1 Fish And Cell Lines

2.1.1a Source And Maintenance Of Channel Catfish Research Specimens

Oculocutaneous albino channel catfish (*Ictalurus punctatus*) were obtained from local pet stores. Although these fish lack pigments in their skin and eyes, they do have pigments in their peritoneal cavities and do have pigmented melanomacrophage cells. Wild-type channel catfish were obtained through Aquatic Imports, Calgary. All fish are believed to be outbred.

The channel catfish were maintained by the University of Alberta Department of Biological Sciences Aquatic Facility staff according to guidelines set forth by the Canadian Council on Animal Care. They were kept at approximately 27°C, an immune permissive temperature [Bly, 1986; Clem, 1984; Lin, 1992], under seasonal light conditions.

All research procedures were pre-approved by the University of Alberta Animal Care and Use Committee and the channel catfish were anaesthetized or euthanized with buffered TMS (50 mg/L Tricaine Methanesulfonate; Syndel Laboratories) prior to all procedures.

2.1.1b Source And Maintenance Of Channel Catfish B-cell Lines

Two immortal B-cell lines derived from channel catfish peripheral blood lymphocytes (PBLs) were used in experiments: 1B10 [Miller, 1994] and 3B11 [Wilson, 1997]. Characterization of these B-cell lines suggests that 1B10 exhibits a phenotype consistent with its having arisen from a previously activated B-cell (IgM+/IgD-) and 3B11 exhibits a naïve mature phenotype (IgM+/IgD+) [Wilson, 1997].

These B-cell lines have been in continuous culture since 1993. Fish B-cell lines were maintained as previously described [Miller, 1984] in Complete Catfish

Medium (cfm) [Miller, 1994] with humidification and 5% CO_2, with the exception that cells were maintained at a slightly higher temperature of 28.5°C. The cfm is composed of equal parts AIM V (Invitrogen) and L-15 (with L-glutamine; Invitrogen) adjusted to channel catfish tonicity (10% dilution with water [van Ginkel, 1994]) and supplemented with 10% fetal bovine serum (FBS; Invitrogen and Sigma; heat-inactivated in-house: 56°C for 30 minutes) and 1% fish serum (carp serum [obtained in-house] or channel catfish serum [a gift from Norman Miller, University of Mississippi Medical Center] heat-inactivated in-house: 50°C for 30 minutes).

2.1.1c Maintenance Of Channel Catfish B-cell Lines Undergoing Experimentation

Antibiotics were not included as a constant constituent of cell maintenance, but were added (100 Units/ml Penicillin, 100 µg/ml Streptomycin, and/or 0.25 µg/ml Amphotericin B; Invitrogen) prophylactically to cultures undergoing manipulations (e.g. FACS) that increased the risk of exposure to bacterial and fungal contaminants.

During experimentation with the cell cultures, extra fish serum (3-5% total) was added to the culture medium. In addition to extra fish serum, the culture medium used for cells undergoing harsh treatment (e.g. electroporation) or growth from low density (e.g. subsequent to FACS) was composed of 25% “cell-conditioned” medium (ccm) [Clem, 1996; Lin, 1992; Miller, 1994]. The ccm is culture medium that has had cells grown in it and thus contains various soluble factors generated by those cells [Freshney, 2000]. It was obtained by the growth of a culture of the specific cells (i.e. the cells of the same type as those to which the conditioned medium would subsequently be added) to log phase, removal of the cells by centrifugation at 400 x g for 10 minutes and then 0.22 µm filter sterilization of the supernatant.

2.1.2 Stimulation Of Fish And Cells To Upregulate Immune Genes To Facilitate AID Cloning And Analysis

2.1.2a Vaccination Of Channel Catfish To Induce An Immune Response

Fish were vaccinated with 0.5 to 1 mg bovine serum albumin (BSA, Initial Fractionation by Cold Alcohol Precipitation, Fraction V, 96-99% Albumin; Sigma) per kilogram of fish by intraperitoneal injection [Vallejo, 1991]. The primary vaccination was a 1:1 emulsion of Freund's Complete Adjuvant (FCA; Sigma) with 2 mg/ml BSA in 0.9x Phosphate Buffered Saline (PBS). Freund's Incomplete Adjuvant (FIA; Sigma) was substituted for FCA in booster vaccinations. At least two booster shots were administered to each fish at four-week or greater intervals [Vallejo, 1991]). Seven channel catfish were vaccinated and those that were maintained for a longer period of time received more boosters (~ 4). Final boosters were given less than three weeks before fish harvest.

2.1.2b Activation Of Channel Catfish B-cell Lines By Stimulation With Mitogens To Upregulate Immune Gene Expression

Stimulation of channel catfish immune cells with mitogens is known to induce in them an activated state characterized by increased cellular proliferation and upregulated expression of immune-related genes [Clem, 1990]. In an attempt to induce AID expression in two channel catfish immortal B-cell lines, they were exposed to a variety of known mitogens and generalized activating compounds.

Approximately 8 million cells at 0.8 million cells/ml, obtained from a culture in log phase, were cultured with various mitogens, individually or in combination: 0.5 mg/ml Lipopolysaccharide (LPS; obtained from *Salmonella typhimurium*; continuous exposure; Sigma) [Lin, 1992; Miller, 1994]; 15% v/v Mouse IgG anti-catfish IgM (αIgM; continuous exposure; obtained in-house from the supernatant of hybridoma 9E1 [Bly, 1986; Sizemore, 1984; van Ginkel, 1994]); 50 ng/ml Phorbol 12-Myristate 13-Acetate (PMA; washed from cells after ~19 hours; Sigma) [Lin, 1992; Miller, 1994]; 500 ng/ml Calcium Ionophore A23187 (CI;

washed from cells after ~19 hours; Sigma) [Lin, 1992; Miller, 1994], 50 µg/ml Concanavalin A (ConA; obtained from Jack Beans – *Canavalia ensiformis*, grade IV; continuous exposure; Sigma) [Bly, 1986; Lin, 1992; Miller, 1994].

This experiment was repeated with culture medium that contained 25% channel catfish immortal T-cell line 28S.3 cell-conditioned medium.

2.1.2c Prolonged Activation Of Channel Catfish B-cell Line 1B10 By Stimulation With Mitogens To Maintain Expression Of The AID Gene In An Upregulated State

Preliminary stimulations indicated, by RT-PCR, that PMA/CI treatment induces AID mRNA expression in the channel catfish 1B10 immortal B-cell line. To maximize AID expression for an extended period of time prior to SHM analysis of the endogenous IgH VDJ exon the cells were repeatedly stimulated by PMA/CI, as above, 12 times on a non-standard interval schedule of approximately 3 to 6 days.

2.2 Extraction, Reverse Transcription, And Sequencing Of Nucleic Acids

2.2.1a Extraction Of DNA From Channel Catfish Cells

The DNeasy Tissue Kit (Qiagen) was used for genomic DNA (gDNA) extraction from channel catfish cell lines according to the manufacturer's protocol for cells grown in suspension.

2.2.1b Extraction Of RNA From Channel Catfish Tissues

Tissue fragments were promptly harvested from euthanized fish, weighed, wrapped in tin foil, frozen in liquid nitrogen, and stored at -80°C until RNA extraction.

Frozen tissues were pulverized with a dry ice-chilled mortar and pestle under conditions that maintained them in a frozen state (i.e. frequent addition of liquid

nitrogen). Powdered tissue was transferred to Trizol (Invitrogen) for total RNA extraction according to the manufacturer's protocol for tissues.

Where indicated, Oligotex mRNA mini kit (Qiagen) was used to further extract poly A+ mRNA from the total RNA sample according to the manufacturer's protocol.

2.2.1c Extraction Of RNA From Channel Catfish Cells

Trizol (Invitrogen) was used for total RNA extraction from channel catfish cell lines according to the manufacturer's protocol for cells grown in suspension.

2.2.1d Reverse Transcription From Channel Catfish mRNA To Generate cDNA

Superscript III Reverse Transcriptase (Invitrogen) was used to reverse transcribe channel catfish mRNA to cDNA with a mixed anchored-oligo dT primer set (P-154; Table 3 in Appendix) or, where indicated, a gene-specific primer according to the manufacturer's protocol but without the addition of an RNase inhibitor.

2.2.1e Extraction Of DNA From Agarose Gel

Subsequent to electrophoresis, DNA was extracted from agarose gel slices with QIAQuick or QIAEX Gel Extraction Kits (Qiagen) according to the manufacturer's protocol or eluted directly into mQH2O at 50°C.

2.2.1f Sequencing Of DNA

To confirm the sequence identities of RT-PCR or PCR products and cloned plasmids, the DYEnamic ET Terminator Cycle Sequencing Kit (Amersham Biosciences) was used to determine the nucleic acid sequence of DNA templates prepared by extraction from agarose gel (described above) or by standard plasmid miniprep [Stowers, 1992] with extraction by chloroform but not phenol. Manufacturer's cycling parameters were followed but reactions were modified

from manufacturer's protocol as follows: 10 µl total reaction volume, 1 µl Sequencing reagent premix, 3 µl DYEnamic ET Terminator Dilution Buffer (Amersham BioSciences).

Following precipitation and drying, sequence reaction products were run on an ABI 3730 sequencer by staff in the Molecular Biology Services Unit of Biological Sciences.

2.3 Cloning Of Channel Catfish AID Gene

2.3.1 Design Of Teleost AID Homologue Degenerate Nested Primers

Due to the lack of available channel catfish sequence, the genome sequences from two proxy teleosts, zebrafish and Fugu, were used as the templates for the design of teleost AID homologue degenerate primers. The National Center for Biotechnology Information (NCBI) zebrafish and Fugu whole genome sequence (WGS) databases were searched with the human AID cDNA sequence (GenBank Accession # AB040431) using tBLASTx, a variant of the Basic Local Alignment Search Tool (BLAST), which compares translated query and database sequences to identify sequences that produce significant alignment. Potential zebrafish and Fugu AID homologue sequences, which were identified in WGS traces since assembled into zebrafish contig 10207.1 (http://www.ensembl.org) and Fugu genomic scaffold 2469 (http://ensembl.fugu-sg.org/Fugu_rubripes), were used to construct predicted zebrafish and Fugu AID homologue cDNA transcript sequences which were then aligned with the AID cDNA transcript sequences of human, mouse (GenBank Accession # AF132979), and chicken (GenBank Accession # AF059262). This alignment of various AID cDNA transcript sequences was used to design a series of degenerate nested primers for use in PCRs intended to amplify teleost AID homologue from channel catfish cDNA (Table 3 in Appendix).

2.3.2 RACE To Obtain Full-Length AID Product

An RT-PCR-amplified fragment of an AID homologue was obtained from the channel catfish 1B10 immortal B-cell line using primers based on conserved regions of putative teleost AID homologues identified from zebrafish and Fugu sequence databases [Saunders, 2004]. This partial sequence was used to inform the design of primers (P-167 and P-170; Table 3 in Appendix) for use with a Rapid Amplification of cDNA Ends (RACE) kit (BD Biosciences Clontech), according to the manufacturer's instructions, to obtain the sequence of the full-length channel catfish AID homologue cDNA from channel catfish spleen.

2.4 AID Expression In Channel Catfish Tissues And Cell Lines

2.4.1 RT-PCR To Examine AID Expression In Select Channel Catfish Tissues

The RT-PCR for channel catfish AID, IgM, and β-actin are fully described in [Saunders, 2004]. Intron-spanning primers were used and the identity of all products was confirmed by sequencing (described above).

For semi-quantitative RT-PCR from cell lines, the mass of RNA added to each RT reaction was equalized.

2.4.2 Northern Blot Hybridization To Examine AID Expression In Select Channel Catfish Tissues

2.4.2a Radioactive Probe Labeling For Northern Blot Hybridization

One channel catfish β-actin and two channel catfish AID intron-spanning regions were chosen as probes for hybridization to a northern blot of channel catfish RNA from selected organs. Primers are detailed in Table 3 (see Appendix) and Probes in Table 1.

Cfβ-actin P-99/P-100: The template for this probe is 474 bp from base 54 to base 527 in GenBank Accession # FD308780.

CfAID P-157/P-161: The template for this probe is 435 bp that covers the 422 bp region from base 193 to base 614 in GenBank Accession # AY436507.

CfAID P-167/P-170: The template for this probe is 422 bp from base 159 to base 580 in GenBank Accession # AY436507.

Radiolabeled probes were generated by Klenow incorporation of [α-^{32}P]dCTP. Probe templates were PCR amplified, electrophoresed on an agarose gel, excised, and gel extracted (described above). A mixture of 5 µl of Klenow Buffer (New England Biolabs), 2 µl of gel extracted probe template, 0.2 µl of Random Hexamer primers, 0.25 µl of each of the two 10 µM specific primers (e.g. P-99 and P-100 for the channel catfish β-actin radiolabeling reaction), and mQH2O to 39 µl was heated to 99°C for 4 minutes and then 2 µl of 10 mg/ml BSA, 2 µl of 500 µM dATP/dTTP/dGTP, 2 µl of Klenow, 5 µl of [α-^{32}P]dCTP were added and the mixture was incubated at 37°C for approximately 90 minutes. After incubation, 0.5 µl of 0.5 M EDTA was added.

Probe was separated from unincorporated nucleotides by centrifugation (1100 x g for 1 minute) of the reaction mixture through a Tris/EDTA equilibrated Select-D (Sephadex G50) column (5 Prime 3 Prime, Inc.).

Probe was boiled for 5 minutes and placed on ice for >1 minute immediately prior to addition to hybridization buffer.

2.4.2b Northern Blot Of mRNA From Select Channel Catfish Tissues

Both immune and non-immune channel catfish tissues were chosen for the examination of AID mRNA expression — intestine, skin, anterior kidney, posterior kidney, spleen, muscle, liver, and brain — to determine if AID is expressed in teleosts in a pattern that is consistent with a role in immune function.

Northern Blot hybridization was as per standard technique. Briefly: Approximately 2 µg mRNA, extracted from total RNA with Oligotex mRNA mini kit (described above), was loaded per sample lane on a 1.5% agarose gel after being heated and then cooled on ice. Ethidium bromide was not added to samples prior to electrophoresis but, subsequent to electrophoresis, the marker

lane containing RNA ladder (Invitrogen) was excised from the gel, ethidium bromide stained separately, and imaged beside a UV fluorescent ruler.

After electrophoresis, mRNA was transferred from the gel to S&S Nytran Supercharge membrane (Scheiler & Schuell) and UV cross-linked in a UV Stratalinker 2400 (Stratagene).

2.4.2c Hybridization Of Channel Catfish AID Radiolabeled Probe To Northern Blot Of mRNA from Channel Catfish Tissues

Hybridization buffer for northern blot hybridization consisted of 50% Formamide, 5x Denhardt's reagent, 0.1% Sodium Dodecyl Sulfate (SDS), 100-200 µg/ml denatured calf thymus DNA, and 5x Sodium Chloride Sodium Phosphate EDTA Buffer (SSPE) in DEPC-treated mQH2O.

Radiolabeled probe generated from CfAID P-167/P-170 template was used to detect channel catfish AID mRNA on the Northern Blot. Pre-hybridization, for 2 hours, and hybridization, for 17 hours, were performed according to Scheiler & Schuell Maximum Strength Nytran membrane with modifications to the washes as follows: 0.5x SSPE/0.1% SDS at 50°C for 20 minutes, 0.4x SSPE/0.1% SDS at 50°C for 20 minutes, 0.4x SSPE/0.1% SDS at 50°C for 20 minutes, 0.4x SSPE/0.2% SDS at 50°C for 20 minutes, 0.4x SSPE/0.2% SDS at 55°C for 20 minutes, 0.2x SSPE/0.1% SDS at 55°C for 20 minutes, 0.1x SSPE/0.1% SDS at 60°C for 20 minutes, 0.1x SSPE/0.1% SDS at 65°C for 75 minutes.

Autoradiograph film was developed after 3 days of exposure and the Northern blot was stripped immediately according to Scheiler & Schuell Maximum Strength Nytran protocol B.

As no signal was detected in the autoradiograph from first AID probe hybridization, the stripped northern blot was re-probed for channel catfish AID mRNA with fresh radiolabeled probe generated from an approximately equal mix of two templates CfAID P-167/P-170 and CfAID P-157/P-161. Hybridization was extended to 19 hours and washes were adjusted as follows: 0.5x SSPE/0.1%

SDS at 50°C for 20 minutes, 0.3x SSPE/0.1% SDS at 60°C for 20 minutes, 0.1x SSPE/0.1% SDS at 65°C for 60 minutes.

Autoradiograph film was developed film after 7 days and the Northern blot was stripped immediately according to Scheiler & Schuell Maximum Strength Nytran protocol B.

2.4.2d Hybridization Of Channel Catfish β-actin Radiolabeled Probe To Northern Blot Of mRNA from Channel Catfish Tissues

Channel catfish β-actin was used as a control for both the presence of mRNA and mass loaded per lane. The twice-stripped blot was hybridized with fresh radiolabeled probe generated from Cfβ-actin P-99/P-100 template.

Hybridization and wash parameters were as for re-probing with CfAID P-167/P-170 and CfAID P-157/P-161.

Autoradiograph film was developed film after 12 hours.

Table 1. Probes for Northern Blot and *In Situ* Hybridization.

Probe Name	**Brief Description**[1]	**Procedure**
Cfβ-actin P-99/P-100	Radiolabeled DNA probe generated with Random Hexamer primers from PCR-amplified channel catfish β-actin cDNA template.	Northern Blot Hybridization
CfAID P-157/P-161	Radiolabeled DNA probe generated with Random Hexamer primers from PCR-amplified channel catfish AID cDNA template.	Northern Blot Hybridization
CfAID P-167/P-170	Radiolabeled DNA probe generated with Random Hexamer primers from PCR-amplified channel catfish AID cDNA template.	Northern Blot Hybridization
IgH	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish IgH cDNA template.	*In Situ* Hybridization
TCRα	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish TCRα cDNA template.	*In Situ* Hybridization
TCRα+3'UTR	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish TCRα cDNA template.	*In Situ* Hybridization
AID	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish AID cDNA template.	*In Situ* Hybridization
AIDN	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish AID cDNA template.	*In Situ* Hybridization
AIDC	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish AID cDNA template.	*In Situ* Hybridization
AID5'3'	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel	*In Situ* Hybridization

	catfish AID cDNA template.	
RAG1	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish RAG1 cDNA template.	*In Situ* Hybridization
TCRβ	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish TCRβ cDNA template.	*In Situ* Hybridization
TCRβ3'	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish TCRβ cDNA template.	*In Situ* Hybridization
IpCD4L1	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish IpCD4L1 cDNA template.	*In Situ* Hybridization
IpCD4L2	DIG-labeled RNA antisense or sense probes generated by *in vitro* transcription from PCR-amplified channel catfish IpCD4L2 cDNA template.	*In Situ* Hybridization

2.4.3 RT-PCR To Examine Mitogen-Induced Gene Expression Changes In Channel Catfish B-cell Lines

Semi-quantitative RT-PCR was chosen to examine mitogen-induced changes in the levels of gene expression of channel catfish AID, IgH, and β-actin in channel catfish 1B10 and 3B11 immortal B-cell lines. Stimulated cells were harvested 3 days after addition of mitogens and RNA was Trizol extracted (described above). RNA was reverse transcribed (described above) and used in PCR reactions as detailed in [Saunders, 2004] for 1B10.

2.5 Sequencing Of Channel Catfish B-cell Line Endogenous IgH

2.5.1 High-Fidelity PCR Amplification Of VDJ From 1B10 Cell Line Stimulated 12 Times With PMA/CI

A high-fidelity polymerase (Pfu, prepared in a departmental fermentation lab) was used in PCR to minimize the introduction of mutation during amplification (PCR artefact) of the VDJ sequence and cycle numbers were kept relatively low. The template was gDNA from 1B10 cells that had been stimulated with PMA/CI 12 times (described above) and primers were P-263/268 (667 bp). The PCR cycling parameters were: 94°C for 3 minutes, 30 cycles of 94°C for 30 seconds and 45°C for 30 seconds and 72°C for 75 seconds, 72°C for 7 minutes.

The PCR products were subcloned (pCR2.1-TOPO, Invitrogen) and sequenced (described above). Sequences were compared with that from an unstimulated control and point mutations were noted.

2.5.2 High-Fidelity PCR Amplification Of VDJ From 1B10 Cell Line Stably Transformed By Tet-Controlled, Ectopic AID Expression System

A high-fidelity polymerase (Phusion, New England Biolabs. Inc.) was used in PCR to minimize the introduction of mutation during amplification of the VDJ sequence and cycle numbers were kept relatively low. The template was gDNA from 1B10 cells that had been stably transfected with Tet-controlled AID gene

expression system (described below) and primers were P-263/268 (667 bp). The PCR cycling parameters were: 98°C for 1 minute, 30 cycles of 98°C for 30 seconds and 57°C for 30 seconds and 72°C for 90 seconds, 72°C for 7 minutes.

The PCR products were subcloned (pCR2.1-TOPO, Invitrogen) and sequenced (described above). Sequences were compared with that from an unstimulated control and point mutations were noted.

2.6 Histology

2.6.1 Preparation Of Channel Catfish Tissues For Histological Examination

Channel catfish were not fed on the day of tissue harvest. Euthanasia was accomplished by an overdose of TMS neutralized to aquarium water pH. The fish were routinely bled with a needle before being fully exsanguinated by a cut to the gill arch and/or caudal peduncle.

Channel catfish tissues were promptly excised and washed and then maintained in cold 0.9x PBS (i.e. channel catfish isotonicity) with 50 Units/ml of heparin (Heparin sodium salt from porcine intestinal mucosa; Sigma) during processing.

Tissues were cut into pieces such that the maximum thickness of the widest dimension was less than 4 mm. These pieces were placed in an excess of freshly prepared 10% neutral buffered formalin (10% NBF = 4% Formaldehyde; Fisher Scientific) in a 15 ml centrifuge tube at room temperature and periodically agitated by hand during the first few hours of fixation. Additionally, tissues from fish vaccinated with fluor-conjugated antigen (describe below) were stored in the dark as much as possible during fixation and all subsequent procedures. The NBF was exchanged at least once, within the first 4 hours of fixation, and more than once if the fixative became turbid (e.g. due to residual mucus).

Fixative was removed after 24-28 hours and tissue pieces were washed with 70% ethanol. Tissues were stored in 70% ethanol for 30 minutes or until paraffin embedding which typically occurred within less than one week after fixation.

Preparation of formalin-fixed tissue for paraffin embedding was based on Standard Fast Procedure [Kiernan, 1999]. Tissues in 70% ethanol were further dehydrated (2 x 30 minutes in absolute ethanol mixed with sufficient Eosin Y to colour tissues enough to enhance visibility during embedding and sectioning) and cleared (2 x 15 minutes in histological grade xylenes; Sigma-Aldrich) before being infiltrated by and embedded in paraffin wax (3 x 30 minutes in Paraplast Plus at less than 65°C; Fisher Scientific).

Blocks of paraffin-embedded tissue were stored in the dark at room temperature and those from fish vaccinated with fluor-conjugated antigen were additionally shrouded in tinfoil.

Formalin-fixed, paraffin-embedded (FFPE) tissues were cut into 5 µm serial sections, floated on warmed DEPC-treated water in a baked pyrex dish, and adhered to Superfrost Plus microscope slides (Fisher Scientific) in a manner consistent with minimizing exposure to RNases.

Slides were stored in opaque microscope slide storage boxes or trays, at room temperature, prior to use.

Tissue sections were cleared of wax by immersion in histological grade xylenes (2 x 10 minutes) and gradually rehydrated through a series of ethanol-to-water incubations (10 minutes each: 1 x absolute ethanol, 1 x 90% ethanol, 1 x 70% ethanol, 1 x 50% ethanol, 2 x DEPC-treated Milli-Q water).

2.6.2 Antigen Transport And Cell Death In Channel Catfish Spleen Sections

2.6.2a Vaccination Of Channel Catfish With Fluor-Conjugated BSA For Confirmation Of Antigen Transport

To confirm that the BSA antigen is transported into the catfish lymphoid tissues, a fish was vaccinated with BSA and given one BSA boost as per BSA vaccination protocol (described above) followed by four boosts consisting of a water-in-oil emulsion, obtained by vortexing, of a 1:1 mixture of FIA (Sigma) and 9 parts of 2 mg/ml Fluorescein-isothiocyanate-conjugated BSA (FITC-BSA; Sigma) + 1 part 2 mg/ml BSA (Sigma) in 0.9x PBS. The final boost with fluor-

conjugated antigen, which should be autonomously identifiable in tissue sections visualized with the appropriate light and filter, occurred two days prior to harvest of fish. Harvested tissues were preserved with 10x FFPE as described above.

A Leica DM RXA compound microscope with a Nikon DXM1200 digital camera using Picture Frame Software was used to visualize fluor-conjugated BSA in dewaxed and rehydrated 5 μm sections of 10x FFPE channel catfish spleen.

2.6.2b TUNEL Cell Death Detection Assay in Channel Catfish Spleen Section

Since the germinal centers of higher vertebrates are sites of extensive cell death [Manser, 2004], the spleens of vaccinated channel catfish were examined for cell death to determine if patterns or discrete regions could be identified and, if so, were they within proximity to regions theorized to be acting in lieu of germinal centers in fish.

The cell death detection assay was performed on 5 μm sections of 10x FFPE channel catfish spleen with the *In Situ* Cell Death Detection Kit, AP (Roche) according to the manufacturer's protocol with the exception that the procedure was halted after primary detection (i.e. incorporation of fluorescein-dUTP).

Cells were counterstained with Hoechst 33342 nuclear stain (Molecular Probes).

2.6.3 In Situ Hybridization

2.6.3a Cloning Of Riboprobe Template Sequences

To identify, *in situ*, the location and organization of the AID-expressing cells in the context of other cells of the humoral immune response, a series of relevant channel catfish marker genes — AID, IgH, TCRα, TCRβ, IpCD4L1, IpCD4L2 — were chosen for use as riboprobes. A RAG1 riboprobe was also chosen to distinguish developing B-cells in the channel catfish kidney which, in addition to its role as a renal and endocrine organ, performs the function of a hematopoietic organ in place of bone marrow and also serves as a secondary lymphoid organ

[Zwollo, 2005]. All riboprobes incorporated one or more intron-spanning regions of each gene.

Channel Catfish IgH Riboprobe Template Cloning

IgH: This covers a portion of the constant region of IgH encoded by Cμ1 to Cμ3 (785 bp from base 520 to base 1304 in GenBank Accession # M27230) and will allow for detection of mRNA transcripts for both membrane-bound and secreted IgM. This was PCR amplified from 1B10 cDNA (see above) with primers P-393 and P-394, digested with restriction endonucleases BamHI (NEB) and XbaI (NEB) and ligated into likewise digested pGEM-3Zf(+) vector (Promega) in forward orientation to generate plasmid pGEM-3Zf-cfIgH.

Channel Catfish TCRalpha Riboprobe Templates Cloning

TCRα: This covers a portion of the constant region encoding sequence of TCRalpha mRNA (354 bp from base 392 to base 745 in GenBank Accession # U58505). This was PCR amplified from channel catfish immortal T-cell line 28S.3 cDNA (reverse transcribed with gene-specific primer P-469) with primers P-467 and P-468 and cloned into pCR2.1-TOPO vector (Invitrogen), according to the manufacturer's protocol, to generate plasmid pCR2.1-TCRalpha. The insert was excised from this plasmid by digestion with SpeI (NEB) and ligated into XbaI digested pGEM-3Zf(+) in forward orientation to generate plasmid pGEM-3Zf-TCRalpha.

TCRα+3'UTR: This covers a portion of the constant region encoding sequence of TCRalpha mRNA (454 bp from base 392 to base 845 in GenBank Accession # U58505). This was PCR amplified from channel catfish immortal T-cell line 28S.3 cDNA (reverse transcribed with gene-specific primer P-469) with primers P-467 and P-469 and cloned into pCR2.1-TOPO vector, according to the manufacturer's protocol, to generate pCR2.1-TCRalpha+3'UTR. The insert was excised from this plasmid by digestion with SpeI and ligated into XbaI digested pGEM-3Zf(+) in forward orientation to generate plasmid pGEM-3Zf-TCRalpha+3'UTR.

Channel Catfish AID Riboprobe Templates Cloning

AID: This covers the entire coding sequence portion of AID mRNA (630 bp from base 52 to base 681 in GenBank Accession # AY436507). This was PCR amplified from pCR2.1-CfAID (a plasmid cloned by Angela Zolner which contains from base 40 to base 758 of channel catfish AID GenBank Accession # AY436507 amplified from channel catfish spleen cDNA with primers P-183 and P-185) with primers P-395 and P-396, digested with restriction endonucleases BamHI and XbaI and ligated into likewise digested pGEM-3Zf(+) vector in forward orientation to generate plasmid pGEM-3Zf-CfAID.

AIDN: This covers the 5'UTR and amino-terminal encoding portion of AID mRNA (395 bp from base 1 to base 395 in GenBank Accession # AY436507). This was PCR amplified from channel catfish spleen cDNA with primers P-573 and P-574 followed by PCR with primers P-579 and P-580 and cloned into pCR2.1-TOPO vector to generate plasmid pCR2.1-CfAIDN.

AIDC: This covers the carboxy-terminal encoding portion and 37 bp of the 3'UTR portion of AID mRNA (236 bp from base 483 to base 718 in GenBank Accession # AY436507). This was PCR amplified from channel catfish spleen cDNA with primers P-573 and P-576 followed by PCR with primers P-575 and P-576 followed by PCR with primers P-579 and P-580 and cloned into pCR2.1-TOPO vector to generate plasmid pCR2.1-CfAIDC.

AID5'3': This covers the portion of AID mRNA sequence covered by both AIDN and AIDC probes (718 bp from base 1 to base 718 in GenBank Accession # AY436507). This was PCR amplified from channel catfish spleen cDNA with primers P-573 and P-576 followed by PCR with primers P-579 and P-580 and cloned into pCR2.1-TOPO vector to generate plasmid pCR2.1-CfAID5'3'.

Channel Catfish RAG1 Riboprobe Template Cloning

RAG1: This covers the amino-terminal half of the coding sequence of RAG1 mRNA (649 bp from base 5 to base 653 in GenBank Accession # AY423858). This was PCR amplified from channel catfish anterior kidney cDNA with primers P-577 and P-578 followed by PCR with primers P-579 and P-580 and cloned into pCR2.1-TOPO vector to generate plasmid pCR2.1-cfRAG1.

Channel Catfish TCRbeta Riboprobe Templates Cloning

TCRβ: This covers a portion of the constant region encoding sequence and 3'UTR in both alleles of TCRβ mRNA (285 bp; see Table 3 in Appendix for GenBank Accession numbers). This was PCR amplified from channel catfish immortal T-cell line 28S.3 cDNA with primers P-566 and P-567 followed by PCR with primers P-579 and P-580 and cloned into pCR2.1-TOPO vector to generate plasmid pCR2.1-cfTCRbeta.

TCRβ3': This covers a portion of constant region encoding sequence and 3'UTR in both alleles of TCRβ mRNA (512 bp; see Table 3 in Appendix for GenBank Accession numbers). This was PCR amplified from channel catfish immortal T-cell line 28S.3 cDNA with primers P-566 and P-568 followed by PCR with primers P-579 and P-580 and cloned into pCR2.1-TOPO vector to generate plasmid pCR2.1-cfTCRbeta3'.

Channel Catfish IpCD4L1 Riboprobe Template Cloning

IpCD4L1: This covers the carboxy-terminal portion of the coding sequence of IpCD4L1 mRNA (600 bp from base 818 to base 1416 in GenBank Accession # DQ435301). This was PCR amplified from channel catfish spleen cDNA with primers P-569 and P-570 followed by PCR with primers P-579 and P-580 and cloned into pCR2.1-TOPO vector to generate plasmid pCR2.1-cfIpCD4L1.

Channel Catfish IpCD4L2 Riboprobe Template Cloning

IpCD4L2: This covers a section of the amino-terminal portion of the coding sequence of IpCD4L2 mRNA (421 bp from base 175 to base 595 in GenBank Accession # DQ435302). This was PCR amplified from channel catfish spleen cDNA with primers P-571 and P-572 followed by PCR with primers P-579 and P-580 and cloned into pCR2.1-TOPO vector to generate plasmid pCR2.1-cfIpCD4L2.

All sub-cloned probe templates were confirmed by sequencing from plasmid preps (described above).

2.6.3b Generation Of Linear DNA Templates *For In Vitro* Transcription Reactions

As linear DNA provides the best template for *in vitro* transcription reactions for riboprobe production, it was generated from plasmid preps of the cloned probe sequences by fully-nested PCR such that the resultant DNA template contained a phage promoter (T7 or SP6) on only one end. Two PCR strategies were employed according to which cloning strategy had been used to generate the specific probe sequence-containing plasmid (i.e. probe sequence in pGEM-3Zf(+) or pCR2.1).

For probe sequence cloned into pGEM-3Zf(+) vector, primers P-472 and P-473 were used to PCR amplify (1 cycle of 94°C for 3 minutes, 5 cycles of 94°C for 30 seconds and 68°C -1°C/cycle for 20 seconds and 72°C for 90 seconds, 35 cycles of 94°C for 30 seconds and 60°C for 20 seconds and 72°C for 90 seconds, and 1 cycle of 72°C for 7 minutes) the region containing the probe sequence and the flanking T7 and SP6 phage RNA polymerase promoter sequences. This PCR product was gel extracted and used in two separate PCR amplification reactions (one with primer P-484 and a gene-specific 3' primer and another with a gene-specific 5' primer and P-485, described below; 1 cycle of 94°C for 3 minutes, 5 cycles of 94°C for 30 seconds and 68°C -1°C/cycle for 20 seconds and 72°C for 90 seconds, 35 cycles of 94°C for 30 seconds and 63°C for 20 seconds and 72°C for 1 minute, and 1 cycle of 72°C for 7 minutes) to generate product with either T7-probe sequence (for *in vitro* transcription of sense strand probe) or probe sequence-SP6 (for *in vitro* transcription of antisense strand probe). PCR products were gel extracted and re-suspended in DEPC-treated mQH2O.

For probe sequence cloned into pCR2.1 vector, primers P-124 and P-466 were used to PCR amplify (1 cycle of 94°C for 3 minutes, 5 cycles of 94°C for 30 seconds and 67°C -1/cycle for 15 seconds and 72°C for 90 seconds, 30 cycles of 94°C for 30 seconds and 62°C for 20 seconds and 72°C for 90 seconds, 1 cycle

of 72°C for 7 minutes) the region containing the T7 phage RNA polymerase promoter sequence-probe sequence-SP6 phage RNA polymerase promoter sequence, but not the vector T7 phage RNA polymerase promoter sequence. This PCR product was gel extracted and used in two separate PCR amplification reactions (one with primer P-579 and a gene-specific 3' primer and another with a gene-specific 5' primer and P-580, described below; 1 cycle of 94°C for 3 minutes, 8 cycles of 94°C for 30 seconds and 65°C -1°C/cycle for 20 seconds and 72°C for 45 seconds, 35 cycles of 94°C for 30 seconds and 57°C for 20 seconds and 72°C for 45 seconds, 1 cycle of 72°C for 7 minutes) to generate product with either T7-probe sequence (for *in vitro* transcription of sense strand probe) or probe sequence-SP6 (for *in vitro* transcription of antisense strand probe). PCR products were gel extracted and re-suspended in DEPC-treated mQH2O.

Channel Catfish IgH Riboprobe Templates

IgH: Amplified from P-472/P-473 PCR product with either P-484/P-394 or P-393/P-485.

Channel Catfish AID Riboprobe Templates

AID: Amplified from P-472/P-473 PCR product with either P-484/P-396 or P-395/P-485.

AIDN: Amplified from P-124/P-466 PCR product with either P-579/P-618 or P-617/P-580.

AIDC: Amplified from P-124/P-466 PCR product with either P-579/P-620 or P-619/P-580.

AID5'3': Amplified from P-124/P-466 PCR product with either P-579/P-620 or P-617/P-580.

Channel Catfish RAG1 Riboprobe Templates

RAG1: Amplified from P-124/P-466 PCR product with either P-579/P-622 or P-621/P-580.

Channel Catfish TCRalpha Riboprobe Templates

TCRα: Amplified from P-472/P-473 PCR product with either P-484/P-468 or P-467/P-485.

TCRα+3'UTR: Amplified from P-472/P-473 PCR product with either P-484/P-469 or P-467/P-485.

Channel Catfish TCRbeta Riboprobe Templates

TCRβ: Amplified from P-124/P-466 PCR product with either P-579/P-611 or P-610/P-580.

TCRβ3': Amplified from P-124/P-466 PCR product with either P-579/P-612 or P-610/P-580.

Channel Catfish IpCD4L1 Riboprobe Templates

IpCD4L1: Amplified from P-124/P-466 PCR product with either P-579/P-614 or P-613/P-580.

Channel Catfish IpCD4L2 Riboprobe Templates

IpCD4L2: Amplified from P-124/P-466 PCR product with either P-579/P-616 or P-615/P-580.

2.6.3c *In Vitro* Transcription To Generate DIG-labeled Riboprobes

Digoxigenin-11-uridine-5'-triphosphate (DIG-11-UTP; Roche) was used to generate DIG-labeled RNA antisense and sense probes through *in vitro* transcription (IVT) with the MAXIscript *In Vitro* Transcription Kit (T7/SP6; Ambion) according to the manufacturer's protocol. The DIG-11-UTP was added to reactions to a final concentration of 0.33 mM and unlabeled UTP was added to a final concentration of 0.17 mM as this ratio was empirically determined to be optimal for *in situ* hybridization (ISH). Reactions were incubated at room temperature for approximately 1 hour. The optional Turbo DNase I treatment step was included, but the reactions were not otherwise purified prior to addition to ISH buffer. Riboprobe was stored at -80°C in 5 µl aliquots to minimize freeze/thaw cycles during use.

A 3 µl sample of each IVT reaction product was visually confirmed and quantified against a standard RNA ladder (3 µl at 1 µg/µl of 0.24 – 9.5 bp RNA

ladder; Invitrogen) by agarose gel electrophoresis and subsequent staining with SYBR Green (20 minutes 1:10,000 in 1x TBE; FMC BioProducts).

The amount of IVT reaction product required for *in vitro* hybridization was empirically determined for each probe.

2.6.3d *In Situ* Hybridization In Channel Catfish Tissues

The channel catfish IgH riboprobe was chosen as a test probe with which to establish empirically optimized ISH parameters based on the protocol of Mitta et al. [Mitta, 2000] because, among the probes to be analyzed, it was anticipated that IgH transcript would be expressed both in a reasonably workable number of cells in the channel catfish spleen tissue and at a relatively high level within those cells.

In situ hybridization was performed on sections of those channel catfish tissues previously determined, from RT-PCR data, to harbour AID-expressing cells - spleen, kidney, intestine, and skin. As a control, ISH was also performed on sections from the channel catfish liver, a tissue that does not appear to harbour AID-expressing cells (as determined by RT-PCR; see Results) but does harbour melanomacrophage cells (i.e. putative channel catfish FDC analogues).

In situ hybridization was done on 5 µm sections of 10x FFPE channel catfish tissues. After dewaxing and rehydration (see above), pre-hybridization of the sections was carried out as follows: a volume of RNase-free hybridization buffer (10x Denhardt's Solution, 4x Sodium Chloride Sodium Citrate Buffer [SSC], 10% Dextran Sulfate [Mandel Scientific Company Inc.], 0.5 mg/ml yeast tRNA [Roche], 0.1 M Dithiothreitol [Mandel Scientific Company Inc.], 0.5 µg/ml Calf Thymus DNA [Sigma], 50% Formamide [Invitrogen], Diethylpyrocarbonate-treated mQH2O to final volume) sufficient to cover the sections was applied to them and the slides were placed in a chamber humidified with 4x SSC/50% Formamide Solution and placed at 55°C for approximately 90 minutes in a Boekel Hybridization Oven.

Riboprobes were thawed and added to fresh hybridization buffer (0.05 µl IgH or 0.5 µl AID IVT reaction per 100 µl of hybridization buffer), placed at 95°C for approximately 3 minutes to denature secondary structures and then placed on ice for 2 minutes before being applied to slides after most of the pre-hybridization buffer had been aspirated off. Slides were returned to ISH oven in their humidified chamber and incubated at 55°C for 30 hours.

After hybridization, slides were flushed with 2x SSC and washed twice for 20 minutes with 2x SSC and once for 20 minutes with 1x SSC on a rotating shaker.

To degrade any residual unbound probe, slides were treated at 37°C for 10 minutes with pre-warmed 37°C 2x SSC to which 20 µg/ml RNase had been added.

Slides were rinsed with 1x PBS and then blocked for 1 hour at room temperature with blocking buffer (1x PBS, 2% Newborn Calf Serum [Invitrogen], 0.2 mg/ml BSA [Sigma]). Blocking buffer was then aspirated off and antibody in blocking buffer was applied as per the probe detection strategy employed (detailed below).

For Detection of Channel Catfish IgH Riboprobes: Sections were incubated for approximately 2.5 hours at room temperature (~ 24°C) and then overnight (~ 24 hours) at 4°C with a 1:1,000 dilution of anti-Digoxigenin Fab fragments to which Alkaline Phosphatase was conjugated (Roche).

Slides were washed twice for 20 minutes with 1x PBS + 0.1% Tween 20 (PBST) and twice for 20 minutes with 1x PBS.

Solutions of Nitroblue Tetrazolium (NBT; Bio-Rad) and 5-Bromo-4-Chloro-3-Indolyl Phosphate (BCIP; Bio-Rad) were prepared according to the manufacturer's protocol. Slides were incubated with Colouration Buffer (0.1 M Tris-HCl, pH 9.5, 0.1 M NaCl, 0.05 M $MgCl_2$) for approximately 5 minutes. Buffer was aspirated off and 4.5 µl NBT and 3.5 µl BCIP per ml of Colouration Buffer and 0.001 M Levamisole were added for a 2-hour incubation and then washed off with PBST to stop the reaction. The slides were held in 1x PBS until imaging.

For Detection of Channel Catfish AID Riboprobes: Sections were incubated approximately 2.5 hours at room temperature (~ 24°C) and then overnight (~ 24 hours) at 4°C with a 1:250 dilution of Mouse IgG anti-Digoxigenin (Roche).

Slides were washed twice for 20 minutes with PBST and twice for 20 minutes with 1x PBS.

Sections were then incubated in the dark for approximately 2.5 hours at room temperature (~ 24°C) and then overnight (~ 24 hours) at 4°C with 1:200 Alexa Fluor 488 Goat IgG anti-Mouse IgG.

Slides were washed three times with 1x PBS on a shaker and held in PBS in the dark until images were captured.

While sections were temporarily mounted in PBS, images were captured on a Leica DM RXA compound microscope with a Nikon DXM1200 digital camera using Picture Frame Software.

After images were captured, slides with NBT/BCIP detection were mounted in 80% glycerol and sealed with Permount (Fisher Scientific) and slides with Alexa Fluor 488 detection were mounted with Fluoromount-G (Southern Biotech).

A series of controls were done with ISH to ensure that probes were hybridizing to specific mRNA and that endogenous enzymes or background levels of non-specific binding were not interfering with the detection of riboprobes, for example, by creating false positives or obscuring legitimate signal. These controls were processed as detailed above, except where indicated, and included: slides processed without riboprobe and without antibody to account for development of chromogen by endogenous phosphatase enzymes; slides processed without riboprobe but with antibody to account for non-specific background binding of antibody or its conjugates; and slides processed with sense-strand riboprobe to ensure probe specificity.

NB: The ISH and detection parameters detailed herein are optimized parameters that were used to obtain the images in the thesis figures. During ISH optimization, ISH was variably replicated with different ISH and detection parameters in tissue sections from three specimens. Although the intensities of

probe and background staining varied with the different ISH and detection parameters, the data were found to be consistent amongst the different specimens. However, due to the malfunction of an external hard-drive that was used for storage of digital microscopy images, the ISH images used in the thesis figures were obtained predominantly from the tissues of one catfish. The ISH and histology images in Figures 7, 12, 13, 14, 15, 16, 18, and 19 were obtained from catfish #1, those in Figure 17 were obtained from catfish #2, and Figure 11 was obtained from a third catfish which had been vaccinated with FITC-BSA.

2.7 Controlled AID Expression In Channel Catfish B-Cell Line

2.7.1a Cloning Of Tet-Off System Plasmids For Controlled Ectopic AID Expression

The Tet-Off system (Clontech) requires the sequential stable integration of two genetic constructs, the pTet-Off control plasmid, which encodes the tetracycline-responsive transcriptional activator (tTA), and the pTRE2 response plasmid, which encodes the AID gene of interest (plasmids are listed in Table 2). The constitutively expressed tTA controls expression from pTRE2 in response to tetracycline or its more stable, higher affinity analogue doxycycline, the use of which is recommended over tetracycline by the manufacturer. The system is repressed (i.e. off) in the presence of tetracycline, as tTA complexed with tetracycline is unable to bind to the pTRE2 response plasmid to initiate transcription.

To facilitate antibiotic selection for Tet-Off expression system transformants in the channel catfish B-cells and be compatible with the selectable markers of existing SHM and CSR reporters, the commercial pTet-Off plasmid (Clontech) was modified by replacement of the Neomycin Phosphotransferase gene, which confers resistance to G418, with the Blasticidin S deaminase gene (in the same orientation as the Neomycin Phosphotransferase gene that it replaced), which confers resistance to Blasticidin, to create the plasmid pTet-Off(Blastres). Digestion with the restriction endonuclease XhoI (NEB) was used to remove the

Neomycin Phosphotransferase gene and it's promoter and polyA signal from the pTet-Off plasmid, which was then treated with Mung Bean Nuclease (NEB) and Calf Intestinal Alkaline Phosphatase (Invitrogen). The Blasticidin resistance gene with its promoter and polyA signal was excised from pcDNA6/TR (Invitrogen) with the restriction endonucleases SmaI (NEB) and SalI (NEB). The fragment was then treated with Mung Bean Nuclease and ligated into the prepared vector.

Likewise, the commercial pTRE2-pur plasmid (Clontech) was modified by replacement of the Puromycin resistance gene, which confers resistance to Puromycin, with the Sh ble gene (in the same orientation as the Puromycin resistance gene had been), which confers resistance to Zeocin, to create the plasmid pTRE2-zeo. Digestion with the restriction endonuclease XhoI (NEB) was used to remove the Puromycin resistance gene and it's promoter and polyA signal from the pTRE2-pur plasmid. The Zeocin resistance gene was PCR amplified with primers P-460 and P-461 from pcDNA4/TO (Invitrogen). The PCR product was sub-cloned into pCR2.1-TOPO and then cleaved out with XhoI restriction endonuclease and ligated into the prepared pTRE2-pur.

Six different AID expression plasmids were prepared by cloning the Kozak and full coding sequences of Catfish AID, Zebrafish AID, Human AID, and the carboxy-truncated sequences of the same into the pTRE2-zeo response plasmid, as described below. In the text, pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc are used to collectively refer to those plasmids with full-length or carboxy-truncated AID sequences, respectively.

Cloning Of pTRE2-zeo-CfAID

Catfish AID was obtained from the plasmid pCR2.1-CfAID (described above). It was digested with the restriction endonucleases BamHI and NotI to excise the P-183/P-185 Catfish AID fragment (719 bp covers base 40 to base 758 in GenBank Accession # AY436507 and includes Kozak sequence and full CDS), which was then ligated into likewise digested pTRE2-zeo.

Cloning Of pTRE2-zeo-ZfAID

Zebrafish AID was PCR amplified from a mixture of zebrafish spleen and kidney cDNA with primers P-391 and P-392 (646 bp covers base 34 to base 679

in GenBank Accession # AY528720 and includes Kozak sequence and full CDS) and cloned into pCR2.1-TOPO vector in forward orientation to generate plasmid pCR2.1-ZfAID. This plasmid was digested with the restriction endonucleases BamHI and NotI to excise the P-391/P-392 Zebrafish AID fragment, which was then ligated into likewise digested pTRE2-zeo.

Cloning Of pTRE2-zeo-HuAID

Human AID was PCR amplified from cDNA generated from Ramos cells, a human B-cell line that constitutively expresses AID, with primers P-207 and P-97 (603 bp covers base 71 to base 673 in GenBank Accession # NM_020661) and cloned into pCR2.1-TOPO vector in forward orientation to generate plasmid pCR2.1-HuAID. This plasmid was digested with the restriction endonucleases BamHI and NotI to excise the P-207/P-97 Human AID fragment, which was then ligated into likewise digested pTRE2-zeo.

Cloning Of pTRE2-zeo-CfAIDCtrunc

Carboxy-truncated Catfish AID was PCR amplified from pCR2.1-CfAID (described above) with primers P-429 and P-430 (597 bp covers 594 bp of channel catfish AID, base 40 to base 633 in GenBank Accession # AY436507 with an extra 3 bp for premature stop codon) and cloned into the pCR2.1-TOPO vector to generate plasmid pCR2.1-CfAIDCtrunc. This plasmid was digested with the restriction endonucleases BamHI and HindIII (NEB) to excise the P-429/P-430 carboxy-truncated Catfish AID fragment, which was then ligated into likewise digested pTRE2-zeo.

Cloning Of pTRE2-zeo-ZfAIDCtrunc

Carboxy-truncated Zebrafish AID was PCR amplified from pCR2.1-ZfAID (described above) with primers P-431 and P-432 (599 bp covers 597 bp of zebrafish AID, base 34 to base 630 in GenBank Accession # AY528720 with an extra 2 bp as a part of the premature stop codon) and cloned into the pCR2.1-TOPO vector to generate plasmid pCR2.1-ZfAIDCtrunc. This plasmid was digested with the restriction endonucleases BamHI and HindIII to excise the P-431/P-432 carboxy-truncated Zebrafish AID fragment, which was then ligated into likewise digested pTRE2-zeo.

Cloning of pTRE2-zeo-HuAIDCtrunc

Carboxy-truncated Human AID was PCR amplified from pCR2.1-HuAID (described above) with primers P-433 and P-434 (556 bp covers 553 bp of human AID, base 70 to base 622 in GenBank Accession # NM_020661 with an extra 3 bp for premature stop codon) and cloned into the pCR2.1-TOPO vector to generate plasmid pCR2.1-HuAIDCtrunc. This plasmid was digested with the restriction endonucleases BamHI and HindIII to excise the P-433/P-434 carboxy-truncated Human AID fragment, which was then ligated into likewise digested pTRE2-zeo.

2.7.1b Electroporation Of Channel Catfish Immortal B-Cells

To generate stable transfectants, cells of the channel catfish 1B10 immortal B-cell line were electroporated in the presence of linearized DNA plasmid for integration and circular carrier DNA plasmid.

Plasmids for transfection were grown in DH5alpha (Invitrogen) or Top10 F' (Invitrogen) cells and were prepared with either Qiagen plasmid prep kits (midi and maxi) or BioCan Scientific Ultraclean Maxiprep kit as per the manufacturer's protocol.

The pBluescript (pBS) plasmid was chosen for use as carrier DNA. The plasmid bearing the genetic elements intended to be randomly integrated into the 1B10 genome was linearized with the restriction endonuclease AhdI (NEB) according to the manufacturer's protocol. Linearized plasmid and carrier plasmid in Tris/EDTA were mixed such that there was 10 µg of each in a total of 20 µl.

Channel catfish 1B10 B-cells from a log phase culture were centrifuged at 400 x g for 10 minutes. The supernatant was aspirated and cells in the pellet were gently resuspended in an excess of fresh serum-free cfm. Centrifugation and supernatant removal were repeated. Cells were resuspended at a concentration of 8 million cells per 180 µl of serum-free cfm.

For each transfection, 180 µl of cell suspension was mixed with 20 µl of the DNA mixture immediately before electroporation and transferred to a 2 mm gap

electroporation cuvette (Molecular BioProducts). Electroporation of cells was performed in with a BTX ECM630 set to 200V, 1200μF, 50Ω.

Subsequent to electroporation, 1 ml of 75% cfm/25% ccm was added to the cells and they were immediately transferred from the electroporation cuvette to a tissue culture plate well containing another 1ml of 75% cfm/25% ccm with 5% fish serum. Cells were allowed a recovery period of 48 hours when electroporated with pTet-Off(Blastres) and 2 hours when electroporated with pTRE2-zeo-XAID or pTRE2-zeo-XAIDCtrunc before the addition of antibiotics for selection (25 μg/ml Blasticidin or 300 μg/ml Zeocin; InvivoGen). Cells were maintained under selection for greater than two weeks with a concentration of antibiotic empirically determined to kill non-transfected cells within two weeks. Un-electroporated 1B10 cells, periodically monitored by staining with the vital dye Trypan Blue (Invitrogen), were also maintained under antibiotic selection alongside the transfected 1B10 cells for comparison to ensure that ample time was allowed for all un-transfected cells to succumb to the selection process.

2.7.1c FACS Sorting of Antibiotic-Selected Stable Transfectants to Establish Clones

Subsequent to antibiotic selection, live cells were FACS sorted one-cell-per-well into 96-well tissue culture plates at the Faculty of Medicine and Dentistry Flow Cytometry Facility. Cells were maintained as indicated above and with selection antibiotic at half the selection concentration. Clones that proliferated were transferred to 24-well tissue culture plates.

2.7.1d Direct-Cell gDNA PCR To Screen Clones For Integration Of Gene Of Interest

To determine if the integration of the complete gene of interest (i.e. tTA or AID) had occurred in antibiotic-selected clones, PCR was performed with nested primers that flank the gene (Table 3 in Appendix).

A volume of cell culture — 50 µl for pTet-Off(Blastres) clones or 100 µl for pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc clones and non-transfected 1B10 — was transferred from the bottom of culture wells, to obtain several thousand cells, to 200 µl thin-walled PCR tubes and centrifuged briefly to pellet cells. The supernatant was removed and the cell pellets were frozen to facilitate cell lysis. Cells were thawed and, to lyse the cells and inactivate enzymes, 19.5 µl of Lysis Buffer (2 µl of 10x PCR Buffer [Invitrogen], 0.2 µl of 1 M β-2-Mercaptoethanol, 2 µl of 20 mg/ml Proteinase K, 15.3 µl mQH2O) was added to the cells and followed with heat treatment at 55°C for 60 minutes and 95°C for 20 minutes.

For the first round of PCR, the outer set of primers was added to the lysed cell mixture in 10 µl of first PCR buffer (1 µl of 10 µM Outer Primer A, 1 µl of 10 µM Outer Primer B, 1 µl of 2.5 mM dNTPs, 1 µl of 10x PCR Buffer, 1.2 µl of 50 mM $MgCl_2$, 0.5 µl of 5 Units/µl Taq polymerase, 4.3 µl of mQH2O). Touch-down PCR (TD-PCR) cycling parameters (1 cycle of 95°C for 3 minutes, 12 cycles of 94°C for 30 seconds and tempX°C -1°C/cycle for 30 seconds and 72°C for 4 minutes, 10 cycles of 94°C for 30 seconds and tempY°C for 30 seconds and 72°C for 4 minutes for pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc or 3 minutes for pTet-Off(Blastres), and 1 cycle of 72°C for 7 minutes) were used during the first PCR.

For the nested PCR with the inner primer set, 1 µl of PCR reaction from the first round of PCR (above) was used as the template. It was added to 24 µl of second PCR buffer (1 µl of 10 µM Inner Primer A, 1 µl of 10 µM Inner Primer B, 1 µl of 2.5 mM dNTPs, 2.5 µl of 10x PCR Buffer, 0.75 µl of 50 mM $MgCl_2$, 0.5 µl of 5 Units/µl Taq polymerase, 17.25 µl of mQH2O). Standard PCR cycling parameters (1 cycle of 95°C for 3 minutes, 25 cycles for pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc or 30 cycles for pTet-Off(Blastres) of 94°C for 30 seconds and tempZ°C for 30 seconds and 72°C for 2 minutes for pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc or 1 minute for pTet-Off(Blastres), and 1 cycle of 72°C for 7 minutes) were used during the second PCR. These PCR products were excised and extracted from agarose gels and sequenced (see above).

pTet-Off(Blastres) Screening PCRs

Outer Primer A and B were: P-335 and P-336.

Inner Primer A and B were: P-219 and P-313.

PCR annealing temperatures were: tempX= 66°C; tempY= 58°C; tempZ= 60°C.

pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc Screening PCRs

Outer Primer A and B were: P-256 and P-257.

Inner Primer A and B were: P-564 and P-565.

PCR annealing temperatures were: tempX= 64°C; tempY= 55°C; tempZ= 52°C.

2.7.1e DNase I Treatment Of Trizol-Extracted Total RNA Samples For Screening Clones For Expression Of Gene Of Interest

Because the gDNA obtained from the cells with the integrated gene of interest provides the same PCR product for those genes as the reverse transcribed mRNA (i.e. cDNA), it was necessary to DNase I treat the Trizol extracted RNA samples to ensure that no contaminating gDNA was present to be amplified during RT-PCR of the gene of interest.

A 20 µl RNA sample was mixed with 10 µl of 10x DNase I Buffer (Ambion recipe: 100 mM Tris pH 7.5, 25 mM $MgCl_2$, 5 mM $CaCl_2$), 0.5 µl of 10 Units/µl DNase I (Roche), and 69.5 µl RNase-free water. The sample was placed at 37°C for 15 minutes and then at 70°C for 5 minutes.

Subsequent to DNase I treatment, the sample was precipitated, washed, resuspended, and spectrophotometrically quantified.

2.7.1f RT-PCR To Screen Confirmed Integrants For Gene of Interest Transcription

pTet-Off(Blastres) Expression

To confirm transcription of tTA from the clones with integrated pTet-Off(Blastres), 2 µl of DNase I–treated RNA (described above) was used in a ¼

volume RT reaction (described above) to generate cDNA for use in PCR. Then 1.5 µl of RT reaction was used as the template in PCR with primers for the Tet repressor (P-221/P-313; Table 3 in Appendix). Standard cycling parameters (1 cycle of 94°C for 3 minutes, 30 cycles of 94°C for 30 seconds and 59°C for 30 seconds and 72°C for 1 minute, and 1 cycle of 72°C for 7 minutes) were used for PCR. To control for RNA degradation and to ensure that amplifications were the result of priming from cDNA and were not due to residual contaminating gDNA in DNase I-treated RNA samples, PCR was also performed concurrently for Cfβ-actin with intron-spanning primers (P-99/P-100) from the same RT reaction and with primers for the Tet repressor (P-221/P-313) with the initial DNase I-treated RNA as the template.

pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc Expression

To confirm repression/derepression of transcription of XAID from the clones with integrated pTRE2-zeo-XAID or pTRE2-zeo-XAIDCtrunc, each clone was divided into two treatment groups and grown for three days with (i.e. repression treatment) and without (i.e. derepression treatment) doxycycline at a concentration of 1 µg/ml. Equal amounts (1250 ng) DNase I-treated RNA (described above) were used in ¼ volume RT reactions (described above) to generate cDNA for use in PCR. Then 2 µl of RT reaction was used as the template in PCR for XAID using primers P-271 and P-565 that amplify from plasmid regions which flank the XAID sequence and so do not amplify endogenously expressed channel catfish AID transcripts. TD-PCR cycling parameters (1 cycle of 94°C for 3 minutes, 8 cycles of 94°C for 30 seconds and 60°C -1°C/cycle for 30 seconds and 72°C for 1 minute, 35 cycles of 94°C for 30 seconds and 52°C for 30 seconds and 72°C for 1 minute, and 1 cycle of 72°C for 7 minutes) were used for PCR.

RT-PCR for Cfβ-actin (P-99/P-100) was done to control for RNA degradation.

PCR was repeated for clones deemed positive for XAID transcription, but the RNA sample was used without RT. This was to ensure that amplification of XAID

results were not due to residual contaminating gDNA in DNase I-treated RNA samples.

The level of endogenous channel catfish AID transcripts in these clones was also examined to ensure that it had not been altered by transfection (e.g. disturbed by exogenous gene insertion) or treatment (e.g. doxycycline in culture medium). A 1 µl volume of RT reaction was used as the template in PCR reactions with a primer set (P-283/P-185) that amplifies endogenous channel catfish AID transcripts, but not the Catfish AID transcripts from the pTRE2-zeo-CfAID and pTRE2-zeo-CfAIDCtrunc. TD-PCR cycling parameters (1 cycle of 94°C for 3 minutes, 8 cycles of 94°C for 30 seconds and 63°C -1°C/cycle for 30 seconds and 72°C for 1 minute, 35 cycles of 94°C for 30 seconds and 55°C for 30 seconds and 72°C for 1 minute, and 1 cycle of °C for 7 minutes) were used.

Table 2. Tet-Off system plasmids.

Plasmid Name	Plasmid Type	Description[1]
pTet-Off(Blastres)	Control	Constitutively expresses the tetracycline-responsive transcriptional activator (tTA).
pTRE2-zeo-CfAID	Response	Expresses channel catfish AID when medium does not contain tetracycline analogues.
pTRE2-zeo-ZfAID	Response	Expresses zebrafish AID when medium does not contain tetracycline analogues.
pTRE2-zeo-HuAID	Response	Expresses human AID when medium does not contain tetracycline analogues.
pTRE2-zeo-CfAIDCtrunc	Response	Expresses carboxy-truncated channel catfish AID when medium does not contain tetracycline analogues.
pTRE2-zeo-ZfAIDCtrunc	Response	Expresses carboxy-truncated zebrafish AID when medium does not contain tetracycline analogues.
pTRE2-zeo-HuAIDCtrunc	Response	Expresses carboxy-truncated human AID when medium does not contain tetracycline analogues.

[1] Refer to text for full description.

3. Results

3.1 Identification Of An AID Homologue Expressed In Channel Catfish

3.1.1 Channel Catfish Express A Moderately Conserved AID Homologue

Translated mouse/human AID sequence was used to obtain predicted fish AID sequences from genomic sequence databases of representative teleost fish species, zebrafish and Fugu, as sufficient genomic sequence data for channel catfish was not available at the time [Saunders, 2004]. The presence of these sequences in the fish databases indicated that fish might possess a gene homologous to that of mammalian AID.

To determine if an AID homologue is expressed by channel catfish, a variety of degenerate DNA primers were designed, based on the alignment of mammalian and predicted fish AID cDNA sequences, for use in semi-nested PCR across potential introns to amplify a fragment of the channel catfish AID transcript sequence from the cDNA of LPS-stimulated channel catfish immortal B-cell line 1B10 [Saunders, 2004]. It is of note that, although numerous combinations of these primers were used in the PCR reactions, only the combinations of primers that prime from sequences located outside of the catalytic domain were productive.

The nucleotide sequence obtained from this initial RT-PCR-amplified fragment of the transcript was used to design primers for use with a Rapid Amplification of cDNA Ends (RACE) kit which allowed for the acquisition of a full-length channel catfish AID cDNA sequence (GenBank Accession # AY436507) from the cDNA of three pooled spleens from non-vaccinated channel catfish [Saunders, 2004]. The full-length channel catfish AID cDNA is 867 nucleotides, of which 51 nucleotides are 5' untranslated region (UTR), 630 nucleotides are coding sequence, and 186 nucleotides are 3' UTR wherein the first base of the canonical polyadenylation signal (AATAAA) resides 19 nucleotides upstream of the cleavage site (Figure 8).

When compared in its entirety, the predicted channel catfish AID protein has approximately 57% amino acid identity and 81% amino acid similarity with the AID protein sequences from either mouse or human (according to protein-protein FASTA algorithm at http://fasta.bioch.virginia.edu/fasta_www2/fasta_list2.shtml; Figure 9). Intriguingly, this is due mainly to good conservation in the amino-terminus, which houses the putative NLS [Ito, 2004; Shinkura, 2004], and the carboxy-terminus, which houses the NES [Ito, 2004] replete with all of the amino acids determined to be necessary for NES function [McBride, 2004], as when only the cytidine deaminase motif (i.e. the putative catalytic active site; Figure 9) of the predicted channel catfish AID protein is compared with the corresponding region from mouse or human AID, the conservation drops considerably to 46% amino acid identity and 63% amino acid similarity, due in part to the presence of eight additional amino acids in the channel catfish sequence in that region.

Although it may be somewhat counterintuitive, as one might reasonably anticipate that the putative catalytic active site (i.e. functional region) would be more highly conserved than the rest of the protein, this lower degree of evolutionary conservation within the putative catalytic active site of the channel catfish predicted AID protein is consistent with that of the predicted AID protein sequences from two other bony fish, zebrafish and Fugu, that also have extra amino acids in this region (Figure 9) [Saunders, 2004]. Given that the most recent common ancestor of channel catfish and zebrafish dates to approximately 130 million years ago and their most recent common ancestor with Fugu dates to approximately 180 million years ago [Santini, 2009], the agreement in this region among the sequences of the three fish indicates that the extra amino acids are likely a consistent feature in most teleosts and not just an isolated peculiarity of channel catfish (i.e. not a recent insertion in the siluriform AID gene).

When compared to the complete predicted AID protein sequences of other teleost fish, obtained from genome sequence databases, channel catfish AID has 78% amino acid identity and 94% amino acid similarity with zebrafish AID and 69% amino acid identity and 89% amino acid similarity with Fugu. Even among these three fish, there is a slight decline in conservation when only the putative

catalytic active site is compared — 69% amino acid identity and 94% amino acid similarity with zebrafish and 65% amino acid identity and 84% amino acid similarity with Fugu (according to protein-protein FASTA algorithm at http://fasta.bioch.virginia.edu/fasta_www2/fasta_list2.shtml).

Importantly, despite the lower degree of evolutionary conservation in the putative catalytic active site — relative to the rest of the protein when compared to mouse or human AID — the channel catfish predicted AID protein was found to have maintained conservation of identity of the amino acids that had been determined to be essential for cytidine deaminase function [Conticello, 2005; Saunders, 2004]. Channel catfish AID also maintains the identity of most amino acids that are known to be required for SHM and CSR mouse/human AID (Figure 9), although this list continues to be refined [Conticello, 2005; Durandy, 2006; Durandy, 2005; Minegishi, 2000; Revy, 2000; Saunders, 2004; Ta, 2003]. Notably, channel catfish AID amino acid positions equivalent to human M6 and A111, which give rise to HIGM2 (no CSR or SHM) when mutated to M6T or A111E [Durandy, 2006], are sites of conservative (channel catfish has leucine) and non-conservative (channel catfish has serine) substitutions, respectively. It is difficult to predict how these specific differences may impact the activity of channel catfish AID. Despite these differences, the otherwise considerable conservation is a strong indication that channel catfish AID may have conservation of function that would be consistent with it mediating SHM in these organisms [Saunders, 2004]. Additionally, the high degree of conservation — likely due in part to the presence of the NES — in the carboxy-terminus, which is required for CSR in higher vertebrates, likewise suggests that channel catfish AID might be able to mediate CSR even though channel catfish IgH genes cannot class switch because they lack the requisite downstream constant regions.

Identification of a channel catfish AID cDNA sequence is an essential step to allow for questions regarding AID expression and biological function in teleosts to be addressed.

*GACTCAGTTTTGCATGACTCAACCACTGAGTTGGTTTCATCTTTCAGAATG***ATG**AGCAAG**

CTGGACAGTGTGCTGCTGACTCAGAGGAAGTTTATTTACCACTATAAGAATGTGCGCTGG

GCTCGTGGGAGGAACGAGACCTACCTCTGTTTTGTGGTCAAGAAACGCAACAGTCCCGAC

TCGCTCTCCTTCGACTTCGGACACCTGCGCAATCGTTCTGGCTGCCATGTGGAGCTTCTC

TTCCTGAGCTATCTTGGGGTACTGTGCCCAGGTTTCTTGGGTTCCGGTGTGGATGGTGTC

AGGGTGGCTTATGCCATCACCTGGTTCTGTTCCTGGTCACCCTGTTCAAACTGTGCCCAT

CGCCTTTCTCGCTTCATGTCTCAGATGCCCAACCTGCGGCTGCGCATCTTCGTCTCGCGC

CTCTACTTCTGTGACGAGGAGGACAGTCAAGAGAGAGAGGGACTCCGTTGCTTGCAGAGG

GCAGGTGTGCAAGTGACAGTCATGACCTATAAAGATTTTTTCTACTGTTGGCAAACCTTT

GTGGCTCAAAATCAGAAGGCTTTCAAGGCTTGGGACGACCTTCACCAGAACTCTATCCGA

CTGTCTCGGAAACTACAGCGAATCCTGCAGCCTAGTGAGTCTGAAGACCTGAGGGATGGC

TTCGCTCTGCTGGGCCTTTAA**TGACATAAGAGTTTACATGACAATCAATTATGGTGCTGT**

TAAAACAAAATTCCTGAGACTTGAATCACTGAAACATTAATAATTTTTTTTTTTCTCCC

TGAATAGCATTATAGCTTTGATAAAGGCGTCTGGGGGCGCTGTGTCCGTGCTACTTATCA

AAATACTACAATAAAGAGTTTTCGTACAAAAAAAAAAAAAAAAAAAA

Figure 8. Nucleotide sequence of channel catfish AID cDNA, GenBank Accession # AY436507.

The full-length channel catfish AID cDNA is 867 nucleotides. The 5' UTR is 51 nucleotides (italicized text), the CDS is 630 nucleotides (bold text), and the 3' UTR is 186 nucleotides (regular text). The canonical polyadenylation signal, which occurs 19 residues upstream of the cleavage site, is underlined. The translation start and stop codons are highlighted by green and red, respectively.

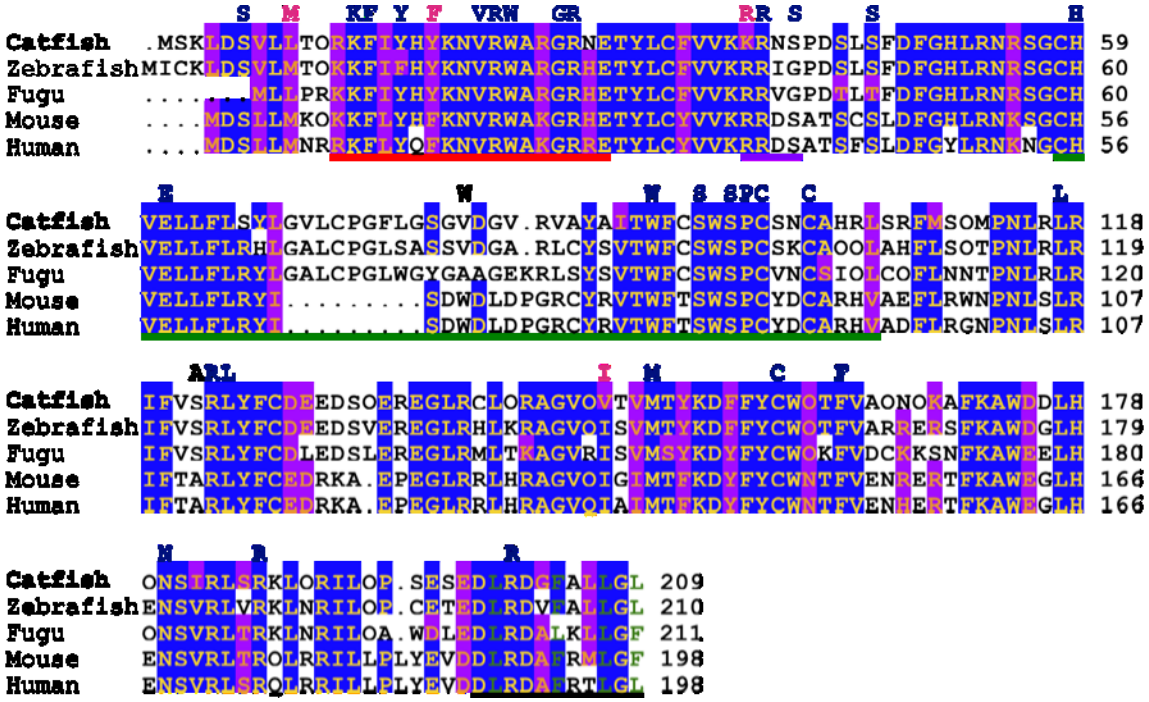


Figure 9. Amino acid sequence conservation among teleost and mammalian AIDs.

Those amino acids determined from mammalian AID studies to be essential to somatic hypermutation, class switch recombination, and/or cytidine deaminase function are indicated above the alignment with their level of conservation in channel catfish AID indicated by blue for identity, pink for similarity, and black for non-conserved. The identity of those amino acids essential for cytidine deaminase catalytic function (H56, E58, C87, C90 in mouse/human AID) has been strictly conserved in the teleost AIDs. And channel catfish AID exhibits conservation with most of the other identified residues. The 11 carboxy-terminal residues, which comprise the nuclear export signal (NES; underlined in black with sites of conserved hydrophobic amino acids necessary for NES function indicated by green text), are also important for class switch recombination function. The putative nuclear localization signal (NLS) is underlined in red, the cytidine deaminase motif in green, and the PKA phosphorylation consensus sequence in purple. AID amino acid sequences are predicted from cDNA (Accession numbers: Catfish AY436507, Zebrafish AY528720, Fugu AY621658,

Mouse AF132979, Human AB040431). Compiled from [Cascalho, 2004; Durandy, 2006; Ichikawa, 2006; McBride, 2004; Minegishi, 2000; Muramatsu, 1999; Revy, 2000; Saunders, 2004; Shinkura, 2004; Ta, 2003; Zhao, 2005].

3.1.2 Channel Catfish AID mRNA Tissue Expression Profile Reveals A Pattern Consistent With A Role In Immune Function

To examine the tissue expression profile of channel catfish AID, intron-spanning primers were used in semi-quantitative RT-PCR with RNA from various lymphoid and non-lymphoid tissues (anterior kidney, posterior kidney, spleen, intestine, fin, liver, muscle, and brain) from exsanguinated channel catfish (Figure 10). These specimens, having been obtained from a local pet store immediately prior to dissection, displayed obvious signs (e.g. minor peripheral injuries accompanied by inflammation) that, although reasonably fit, they were likely engaged in ongoing adaptive immune responses (i.e. a state in which AID would be expected to be expressed) and so did not require vaccination to raise an immune response.

Channel catfish AID mRNA transcripts were visibly amplified from both primary and secondary lymphoid tissues — anterior kidney, posterior kidney, spleen, intestine — and fin but only after a relatively high number of PCR cycles (Figure 10). In aliquots removed after 20 cycles of PCR, there was no visible AID product while the IgH product was visible in the anterior and posterior kidney samples — tissues that are areas of B-cell development and maturation and so could be expected to harbour more IgH-expressing cells than the other tissues examined. To obtain the visible AID products in Figure 10 required 30 cycles of PCR, a number that resulted in amplification of IgH product from all examined tissues except brain (not shown).

While this method revealed that the tissue expression pattern of AID mRNA in channel catfish is consistent with it having a role in immune function, the relatively high number of PCR cycles required to obtain visible AID product suggests that it is expressed at relatively low levels in the channel catfish tissues examined. This supposition is further supported by the inability, despite two attempts with different probe mixes, to detect hybridization of radiolabeled channel catfish AID probes with channel catfish AID mRNA on a Northern blot of mRNA from various channel catfish tissues, even though a radiolabeled channel

catfish β-actin probe, subsequently hybridized to this same blot, produced strong signal (not shown).

Although hybridization problems exclusive to the radiolabeled AID probes cannot be entirely discounted, when corroborating data from the semi-quantitative RT-PCR is taken into consideration, it seems probable that the inability to detect any signal from the AID probe hybridizations was due to insufficient sensitivity of the method to detect the meager amount of AID mRNA present in the tissue samples. In fact, a low level of AID mRNA expression in tissues, including secondary lymphoid tissues, is consistent with mammalian data [Muramatsu, 1999; Muto, 2000; Saunders, 2004].

To better refine regions of AID mRNA expression within the tissues, ISH, a more sensitive method of detecting AID mRNA, was chosen to pursue cellular expression analyses in histological sections from the channel catfish tissues that were found to express AID in the semi-quantitative RT-PCR. In concert with other techniques to establish the environment of AID-expressing cells, this approach should also allow for the resolution of questions about potential germinal center analogue architecture in teleosts.

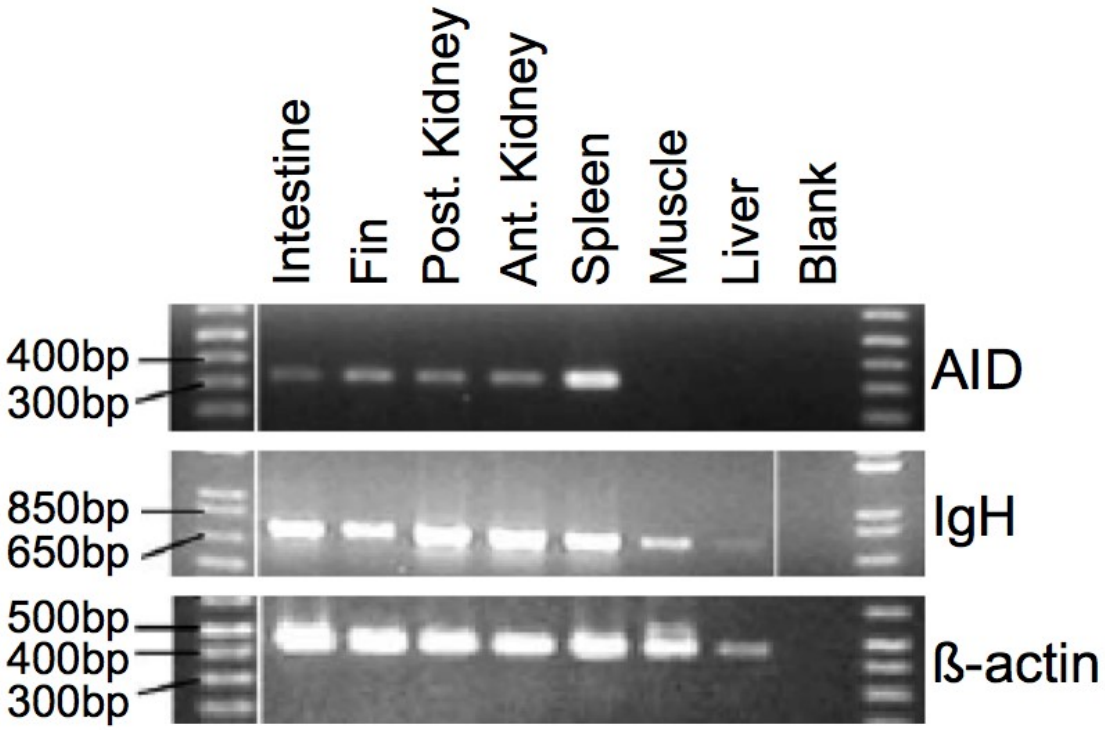


Figure 10. AID mRNA is expressed in a subset of channel catfish tissues.
AID mRNA expression profile of non-lymphoid and lymphoid tissues in the channel catfish indicates that AID is expressed in both primary and secondary lymphoid tissues. Note that in teleost fish the kidney serves as the hematopoietic organ. Comparison of RT-PCR for AID with that for IgH shows that AID is expressed in tissues that harbour plasma B-cells, which are expected to occur in the vicinity of activated B-cells expressing AID, but that not all tissues that harbour plasma B-cells have AID expression. Fish skin (i.e. fin) is a mucosal epithelial tissue and as such may have some secondary lymphoid function. After 20 PCR cycles, aliquots removed from AID PCR had no visible product and those from IgH PCR had visible product only in the anterior and posterior kidney tissues. It was necessary to use 30 PCR cycles to obtain the product pictured here. Identity of all products was confirmed by sequencing. This experiment was replicated with fish sourced from a different supplier and separate equipment. Modified from [Saunders, 2004].

3.2 Identification Of Antigen And Apoptotic Cells In Channel Catfish Spleen

3.2.1 FITC-BSA Was Transported From The Vaccine Injection Site In The Peritoneal Cavity To The Spleen In Channel Catfish

In sections of FFPE spleen tissue obtained from a channel catfish that had been injected with FITC-BSA two days prior to tissue harvest, FITC-fluorescence is conspicuously visible in or on some individual cells (Figure 11), which confirms that the antigen was transported into the spleen from the site of injection in the peritoneal cavity. Spleen sections from catfish vaccinated with unlabeled-BSA did not exhibit this fluorescence.

Cells with FITC fluorescence were scarce in the sections. They did not appear to be either clustered together or colocalized with melanomacrophage cells nor did the melanomacrophage cells themselves appear to bear FITC fluorescence. However, the possibility that there was further but unobservable fluor-conjugated antigen present in or on more of the cells (e.g. trapped on the surface of the melanomacrophage cells) cannot be excluded as the fluorescence capability of the FITC may have been inactivated (e.g. adulterated by binding, digestion, or pH) or inadequate (e.g. low concentration and/or masked by autofluorescence of pigments in melanomacrophage cells [discussed below] [Saunders, 2010]) for visualization under the circumstances.

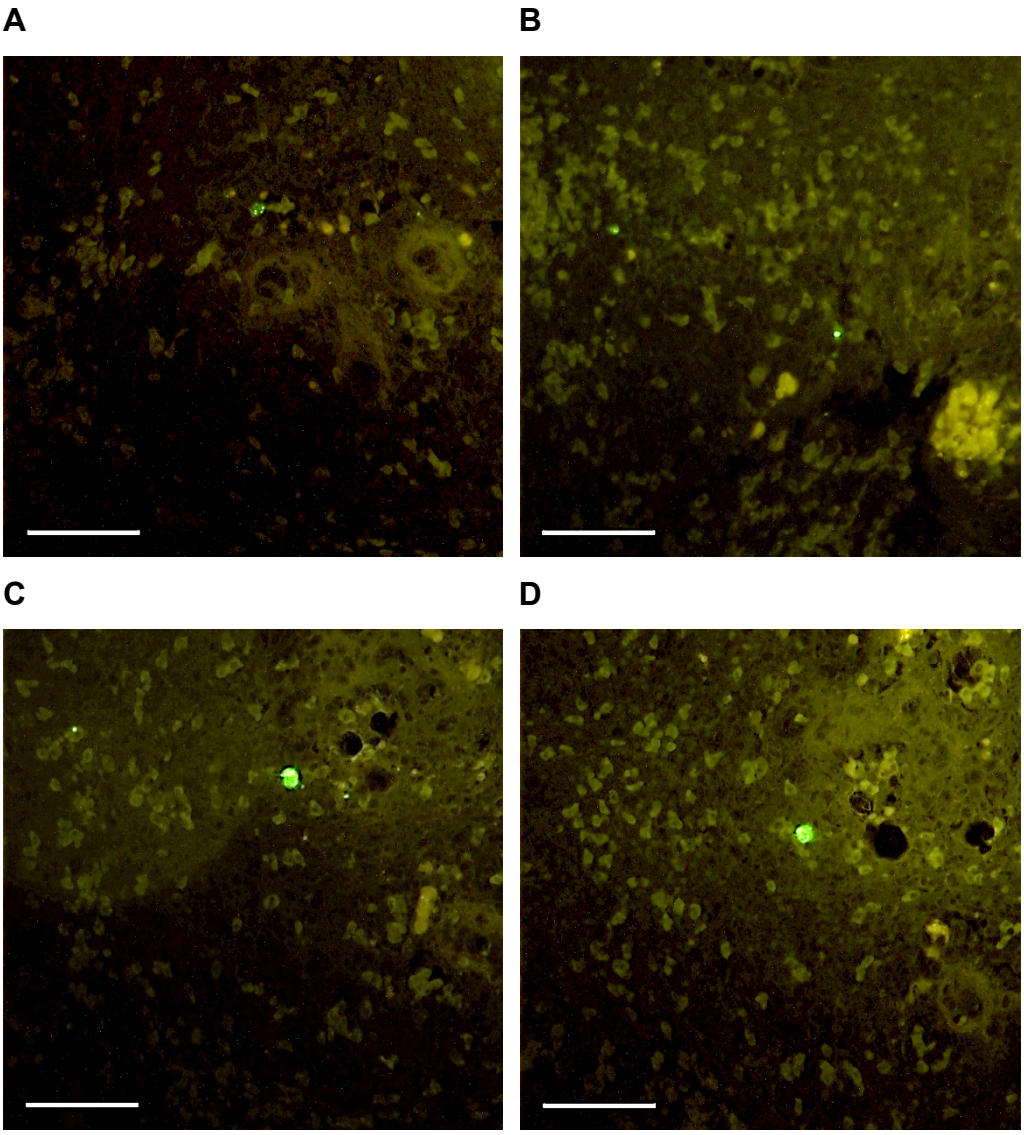


Figure 11. Fluorescence from FITC-BSA can be detected in channel catfish spleen.

The appearance of FITC fluorescence (bright green) in channel catfish spleen sections demonstrates that FITC-BSA (i.e. antigen) was transported into the spleen from the site of vaccine injection in the peritoneal cavity. Comparable green fluorescence was not observed in spleen sections from catfish vaccinated with unlabeled-BSA. Autofluorescent melanomacrophage cells (yellowish) can also be observed in these images which were captured under the green fluorescence channel. Image parameters (i.e. brightness, contrast, and

sharpness) were adjusted post-capture to improve visibility of background cells to provide context for FITC-bearing cells. Scale bar represents 50 μm.

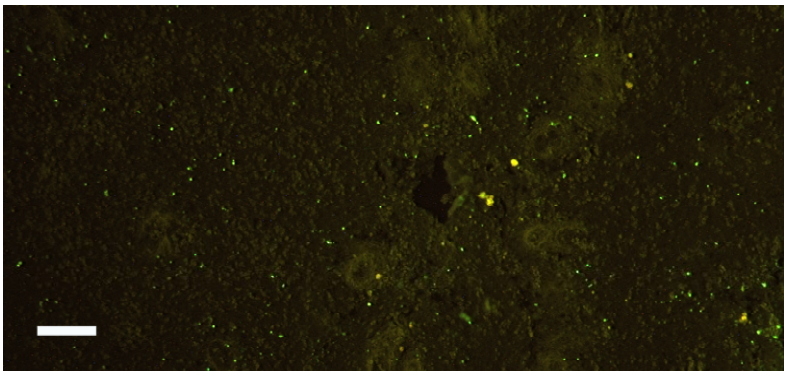
3.2.2 Channel Catfish Spleen Harbours Apoptotic Cells But Their Arrangement Does Not Resemble The Pattern Observed In Mammalian Germinal Centers

The presence of relatively high numbers of apoptotic cells is a morphological signature of germinal centers in mammals. These clusters of apoptotic cells are composed predominantly of non-selected B-cells (i.e. those B-cells that have not received survival signals during antibody affinity maturation). The Terminal deoxynucleotidyl transferase-mediated dUTP Nick-End Labeling (TUNEL) procedure, a reliable assay for identifying *in situ* cell death including that from apoptosis, was performed in channel catfish spleen sections to search for apoptotic cells, especially those occurring in clusters, as indicators of histological sites of potential germinal center analogues. It is important to note that no firm conclusions can be drawn as to the identities of cells stained solely by TUNEL or their precise mode of death as TUNEL merely allows detection of DNA fragmentation by labeling 3'-OH termini, although the manufacturer's manual (Roche) does indicate that an abundance of the favoured substrate of the TUNEL reaction is most likely to be generated by apoptosis. Because these tissues were obtained from relatively healthy fish (i.e. no necrosis in organs), TUNEL-positive cells will be presumed to be apoptotic cells.

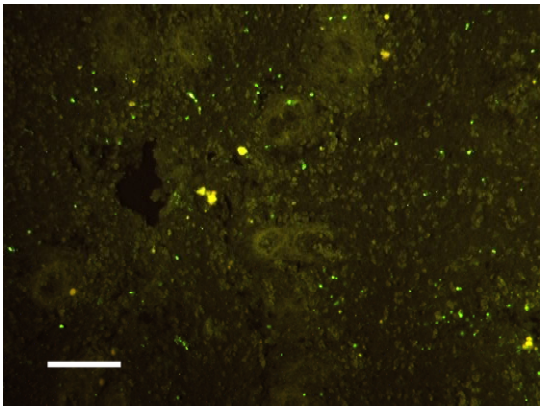
Numerous apoptotic cells were detected in the channel catfish spleen (Figure 12), but their organization did not approximate the clustered pattern observed in antibody affinity maturation in higher vertebrates. Apoptotic cells were distributed diffusely throughout the spleen sections, in both the red and white pulp (see Figure 14 for splenic regions), with no discernible organization. There were no apparent clusters of apoptotic cells, even in the vicinity of melanomacrophage centers as might be expected if these regions serve as primitive germinal center analogues. Further to the lack of clustering, there was no discernible positive association of the apoptotic cells with the melanomacrophage centers. But it cannot be ruled out that the strong autofluoresce of the melanomacrophage cells may have masked the presence of TUNEL-labeled FITC cells (discussed below) [Saunders, 2010].

A similar scattered pattern of apoptotic cells was also observed in the anterior kidney (not shown), which, as a site of considerable apoptosis owing to its role as a hematopoietic tissue, was examined concurrently as a control tissue for the assay and a comparator for the spleen.

A



B



C

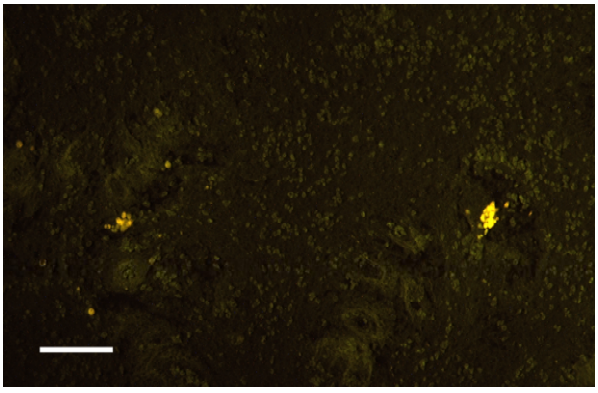


Figure 12. Apoptotic cells are scattered throughout the channel catfish spleen.

Cells containing significant amounts of fragmented DNA (i.e. apoptotic cells), fluorescein-labeled via TUNEL, are scattered throughout both the red pulp and

the white pulp of channel catfish spleen (A). Melanomacrophage cells (yellow) do not appear to have clusters of apoptotic cells associated with them (B, enlarged from image A). Terminal deoxynucleotidyl transferase has been omitted from the negative control (C). Performed in duplicate from one catfish. Scale bars represent 50 μm.

3.3 In Situ Identification Of AID-Expressing Cells In Channel Catfish Spleen

3.3.1 ISH Revealed Some Histological Context Of AID-Positive Cells

Once the channel catfish tissues that harbour AID-expressing cells had been identified by RT-PCR, it was then necessary to examine those tissues, the chief amongst them being the spleen, to identify, *in situ*, the location and organization of the cells that may be involved in AID-mediated antibody affinity maturation processes in fish and so might denote a cellular architecture that serves them in lieu of the germinal centers found in higher vertebrates. To this end, a selection of relevant channel catfish genes was chosen for use as ISH riboprobes to establish the local context of AID-positive cells in the spleen and various other tissues, which were chosen for control and comparison with the spleen. These genes, from which one or more intron-spanning regions of each were chosen for probe sequence, included AID, IgH, TCRα, TCRβ, IpCD4L1, IpCD4L2, and RAG1. Additionally, because it had been hypothesized that melanomacrophage centers might contribute to teleost antibody affinity maturation (see Introduction), the relationship of ISH stained cells with melanomacrophage centers was also assessed.

The AID gene was chosen as a marker with which to specifically identify teleost equivalents of mammalian germinal center B-cells (i.e. B-cells undergoing antibody affinity maturation) as indicators of potential germinal center analogues. As it was expected that AID mRNA transcript levels per cell would be relatively low, a selection of AID riboprobes of different sizes and sequences were designed to better ensure that one could be found empirically that would perform well (i.e. generate ample signal with the detection method employed) and sufficiently overcome issues that often plague ISH (e.g. adequately penetrate the tissue).

The IgH gene was chosen as a marker with which to broadly identify all B-cells. An IgH riboprobe was designed that allowed for the detection of the mRNA transcripts for both the membrane-bound and secreted forms of IgM. Once

identified, the location and organization of these IgH-positive B-cells could be compared and contrasted with the location and organization of the subset of B-cells that were AID-positive.

In addition to B-cell detection, because the IgH-expressing B-cells were expected both to be relatively numerous, especially within the spleen, and to express abundant levels of IgH mRNA transcript per cell, this riboprobe was determined to be the most suitable candidate for use in the initial optimization of the ISH protocol within the fixed channel catfish tissue sections.

The TCRα, TCRβ, IpCD4L1, IpCD4L2 genes were chosen as markers with which to somewhat broadly identify helper T-cells. In higher vertebrates, helper T-cell interactions with germinal center B-cells are an essential component of antibody affinity maturation and these cells are found interspersed with the B-cells and FDCs in the light zone and clustered around the germinal center in the T-cell zone. Identification of T-cells, especially in relatively denser groups, in proximity to or interspersed with AID-positive cells would serve to bolster claims of primitive germinal center-like arrangements in channel catfish.

In fish, which lack bone marrow, the kidney, in addition to its role as a renal and endocrine organ, performs the function of a hematopoietic organ in place of bone marrow and also serves as a secondary lymphoid organ [Zwollo, 2005]. Even though, at least in the trout kidney, hematopoiesis occurs mainly in the anterior portion of the organ and the secondary lymphoid function occurs mainly in the posterior portion of the organ [Zwollo, 2005], semi-quantitative RT-PCR for AID from channel catfish anterior and posterior kidney samples generated similar amounts of product (Figure 10)[Saunders, 2004]. This created uncertainty as to whether, in addition to secondary modification during antibody affinity maturation, channel catfish AID might be involved in primary repertoire development, as occurs in some mammals (see Discussion). Because RAG1, which fulfills the same recombinase role in fish as in higher vertebrates, is expressed in developing B-cells, the RAG1 gene, which has an intron in channel catfish, was chosen as a marker with which to specifically identify developing B-cells to

attempt to determine where these developing cells occur in relation to AID-positive B-cells in the channel catfish kidney.

As detailed below, ISH detection of IgH- and AID-expressing cells was successful in various tissues but the targets of the other tested riboprobes could not be detected in any of the tissues tested. Additionally, attempts to detect AID-expressing cells in skin sections were impeded by technical difficulties that appeared to be specific to the skin sections coupled with the detection method employed (i.e. intense background staining with Alexa Fluor 488).

3.3.2 Melanomacrophage Cells' Autofluorescence Hinders Signal Detection

The pigments of the melanomacrophage cells autofluoresce strongly, which makes it easy to identify the cells *in situ* without the need for any sort of differential staining (Figure 7) [Saunders, 2010]. While this can facilitate the assessment of the relationship of melanomacrophage centers with cells labeled by ISH or other techniques, it can also obfuscate assessment by masking the signals [Saunders, 2010]. This masking effect can be seen in the image of a spleen section stained with Hoechst nuclear stain where the blue-fluorescing nuclei are visible in cells throughout the section except in the melanomacrophage centers (Figure13) [Saunders, 2010].

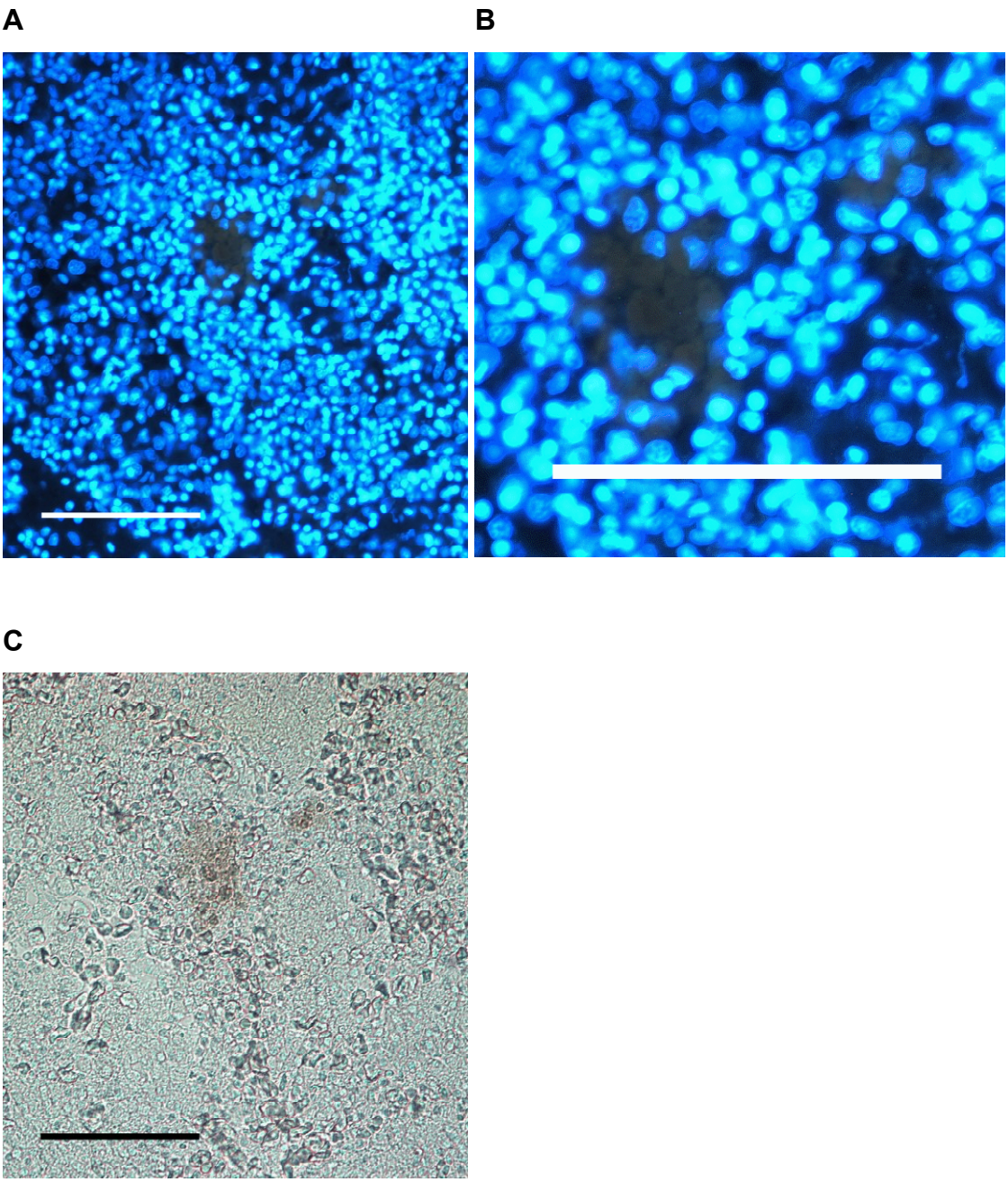


Figure 13. Pigments in melanomacrophage cells mask detection of signals. In this section from channel catfish spleen, the Hoechst-stained nuclei of the cells fluoresce intensely blue under the UV channel (A) but no Hoechst fluorescence is observed in the melanomacrophage cells (B, an enlargement from A), which autofluoresce red in this image. These melanomacrophage centers can also be

observed (brown) in bright field of the same section (C). Scale bars represent 50 μm.

3.3.3 IgH-Positive Cells Do Not Appear To Preferentially Associate With Melanomacrophage Cells

The ISH with NBT/BCIP for detection of IgH-positive cells in tissue sections initially appeared to have been very successful, likely due to the relatively large number of IgH mRNA transcripts present in each B-cell. However, it was later determined that the detection method employed (i.e. NBT/BCIP) was not sufficiently sensitive to detect relatively rare transcripts (see below). Thus, if B-cells with lower expression of IgH were present, as has been proposed for centroblasts in mammals (see Introduction), they were likely not detected — especially if they are associated with the pigmented melanomacrophage centers, which obscure signal (Figure 13). Overall, enzymatic staining of IgH-positive cells with NBT/BCIP chromogen was intense and definite, but nominal background staining did result in a pinkish cast in some sections — especially those prepared during optimization of ISH parameters. Additionally, it was noted that melanomacrophage centers appeared slightly chromogenically enhanced under bright field illumination in sections exposed to NBT/BCIP staining. This was investigated and determined to be a probe-independent side effect of NBT/BCIP staining that could not be entirely eliminated — even by the inclusion of an inhibitor of endogenous alkaline phosphatases (Levamisole) during incubation with the chromogenic substrate mix. This may have further complicated detection of any B-cells with lower expression of IgH (see below).

Relative numbers of IgH-positive cells among the various tissues were found to be generally consistent with IgH RT-PCR data.

In the channel catfish spleen from ISH with NBT/BCIP, the numerous IgH-positive cells appear to be somewhat randomly distributed, almost entirely within the white pulp (Figure 14) [Saunders, 2010]. Some IgH-positive cells do occur in contact with or quite proximate to a few others giving the appearance of small loosely associated clusters (compare this to the large densely associated clusters seen in the anterior kidney described below; Figure15). But it is not clear from these sections whether any or all of these apparent small clusters of IgH-

positive cells represent real interactions (e.g. actively clustered) or are merely a random result of IgH-positive cell density within the limited confines of the white pulp regions. As in the mouse and human, the white pulp of the channel catfish spleen is a dense lymphocyte-rich region that surrounds the blood vessels (Figure 14) [Fange, 1985; Grizzle, 1976; Saunders, 2010]. Splenic white pulp houses germinal centers in the mouse and human and melanomacrophage centers in the channel catfish. The channel catfish spleen exhibits less distinct separation between splenic white pulp and the looser red pulp, which is rich in erythrocytes, than is found in the spleen of mouse and human.

While ISH with NBT/BCIP revealed that some IgH-positive cells in the spleen were located within or at the periphery of melanomacrophage centers, many had no apparent association with melanomacrophage centers. This is also true of the possible small clusters of IgH-positive cells. Whether those IgH-positive cells located proximate to melanomacrophage centers represent differential or preferential interactions by subsets of B-cells or are merely a coincidence that arises from both B-cell and melanomacrophage cell types occupying the same sub-region of the tissue (i.e. white pulp) could not be determined from IgH mRNA staining alone.

Although ISH with NBT/BCIP did not reveal very many IgH-positive cells in melanomacrophage centers, the later discovery — using ISH with detection via fluorochrome-labeled secondary antibody (fluorochrome dye Alexa Fluor 488) — that AID-positive cells associate with melanomacrophage cells (see next section) indicated that substantial numbers of B-cells are present in some melanomacrophage centers (compare Figures 14 and 17). And using melanomacrophage centers obtained by laser-capture microdissection as the template for IgH RT-PCR [Oko, 2009] revealed that IgH is expressed in melanomacrophage centers [Saunders, 2010]. This prompted another attempt to identify IgH-positive cells in the melanomacrophage centers of channel catfish spleen sections using ISH. The ISH protocol was altered to incorporate probe detection via FITC-conjugated secondary antibody — instead of the combination of alkaline phosphatase-conjugated anti-DIG Fab fragments and NBT/BCIP — to

eliminate the probe-independent deposition of coloured precipitate within the melanomacrophage centers. And confocal laser scanning microscopy, a more sensitive imaging apparatus (e.g. greater resolution), was used for image capture. This allowed the epifluorescence of the FITC ISH signals (i.e. IgH-positive cells) to be unmixed from the autofluorescence of the melanomacrophage cells [Saunders, 2010]. With this refined approach, IgH-positive cells were identified as distinct entities within the melanomacrophage centers [Saunders, 2010]. Thus, although IgH-positive cells do not appear to preferentially associate with melanomacrophage cells, those IgH-positive cells that do occur in association with melanomacrophage cells appear to represent a subset of IgH-positive cells that express lower levels of IgH mRNA (see Discussion).

Note that this refined detection and imaging method was not used with other tissues to reexamine melanomacrophage centers for IgH-positive cells or to resolve other signal related issues (e.g. TUNEL) due to time constraints (see Discussion). Therefore, references to IgH-positive cells in tissues other than spleen are based solely on the analysis of less-sensitive NBT/BCIP ISH and so may not include B-cells that express low levels of IgH mRNA.

In the channel catfish kidney, the numerous IgH-positive cells appear to be restricted to and/or excluded from certain regions within the tissue (Figure 15). As the kidney is a multifunctional organ [Zwollo, 2005], the apparent restriction likely results from spatial partitioning of renal, endocrine, and lymphohematopoietic subfunctions within the organ.

In both the anterior and posterior kidney, the IgH-positive cells appear to occur at high density, especially in the anterior kidney where large dense clusters are also a common feature. Given that the trout kidney exhibits a lymphohematopoietic anterior-posterior functional gradient, wherein B-cell development predominates in the anterior (i.e. primary lymphoid) and B-cell maturation predominates in the posterior (i.e. secondary lymphoid), that significant clusters of IgH-positive cells can be found in the channel catfish anterior kidney but not in the posterior kidney or in the spleen, another secondary

lymphoid tissue, suggests that the clustering of IgH-positive cells in the anterior kidney is probably associated with B-cell development processes rather than B-cell maturation processes.

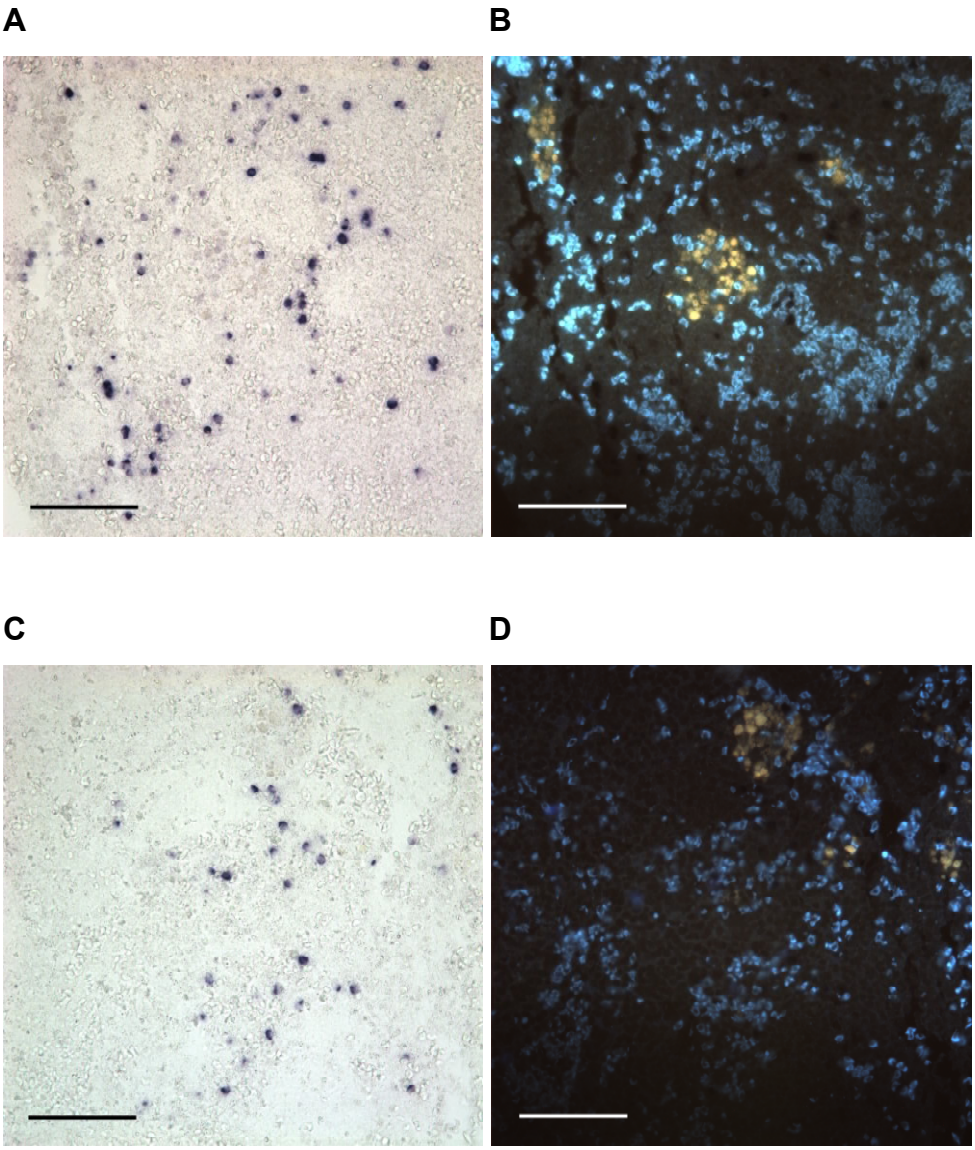
Some of the IgH-positive cells and clusters in the kidney colocalize with melanomacrophage centers while others do not. Thus, as in the spleen, it is difficult to determine if these are positive associations or merely coincidence.

In the channel catfish posterior intestine, a mucosal epithelial tissue, the IgH-positive cells are found in the mucosa where they occur along the basal surface of the lamina propria immediately adjacent to the muscularis mucosae (Figure 16). They appear to be tightly associated with this juncture, occurring in an almost single layer along it. This is distinctly different from the channel catfish skin, another mucosal epithelial tissue, in which the IgH-positive cells are distributed throughout the mucosa and tend to occur more toward the apical surface and less often near the basal surface (Figure 16). As in spleen and kidney, there was no discernible pattern of association of IgH-positive cells with melanomacrophage centers, which in these mucosal tissues tend to be smaller and appear to be more loosely associated than those found in spleen and kidney.

It is of note that, for practical reasons, the skin samples for RT-PCR and histology were obtained from different locations on the fish body. While the skin sample for RT-PCR was obtained from fin to minimize contamination with underlying muscle, the skin tissue for histology was obtained instead from the main body of the fish to allow for improved fixation and easier sectioning. While muscle samples per se were not investigated with ISH, its presence in the skin sections allowed for cursory analysis. IgH-positive cells were rare in muscle immediately underlying the skin and melanomacrophage cells were not observed (not shown).

In the channel catfish liver, no IgH-positive cells were identified. This is considered to be a bona fide result rather than a failure of the ISH, which was very effective for detection of the plentiful IgH mRNA transcripts, because the liver, as neither a primary or secondary lymphoid tissue, was not expected to harbour populations of IgH-positive cells. Furthermore, the apparent absence of

IgH-positive cells in the liver is consistent with RT-PCR data (Figure 10) wherein a relatively large number of cycles were required to amplify IgH from the tissue. However, the liver does have melanomacrophage centers comprised of a few to several cells (Figure 17).


A
B
C
D

E

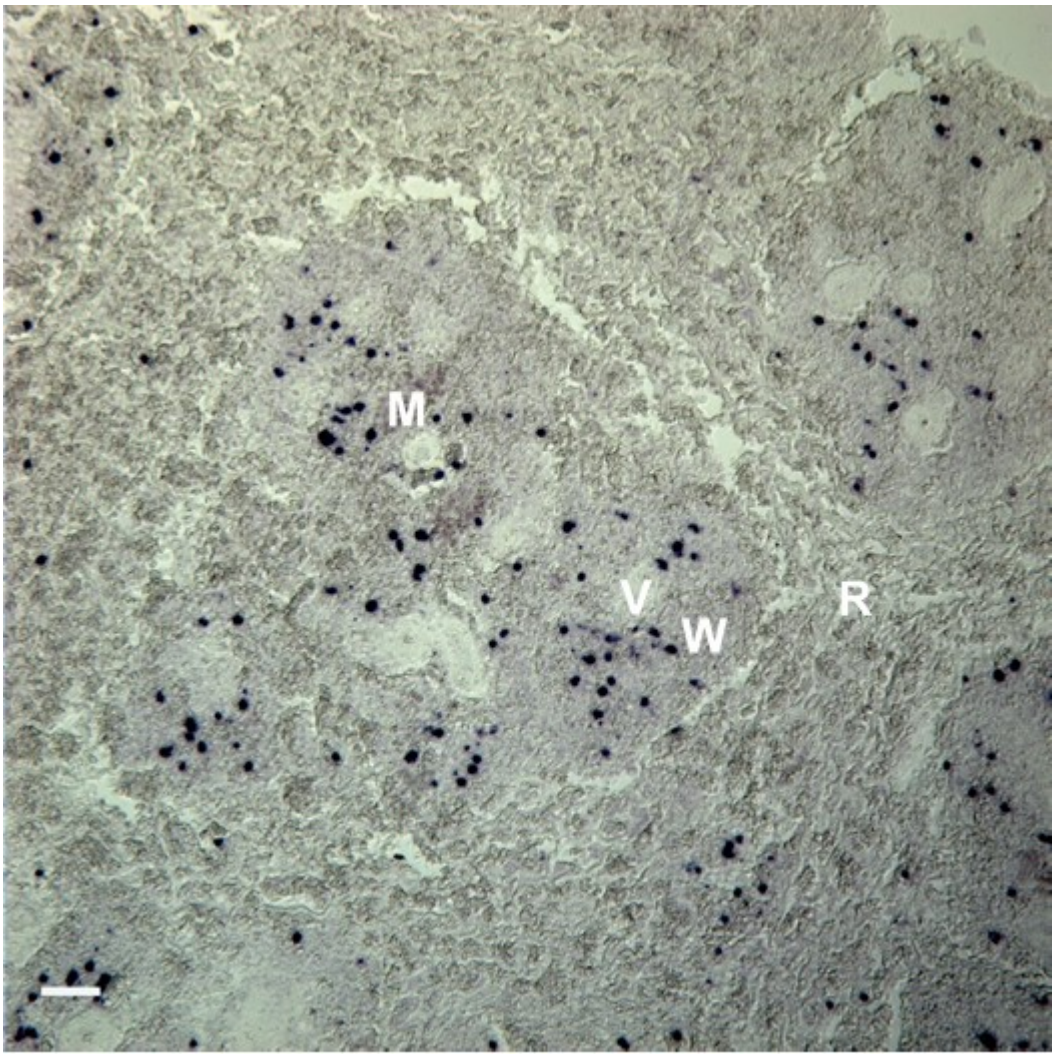


Figure 14. IgH-positive cells occur mainly in the white pulp in channel catfish spleen.

In channel catfish spleen sections, cells chromogenically-labeled by ISH for IgH mRNA (both membrane-bound and secreted IgM) appear dark blue under bright field microscopy (A and C). Although the IgH-positive cells appear to be distributed mainly within the white pulp (A and C), where the melanomacrophage centers (autofluoresce yellow under UV channel, B and D) also occur, no specific association is apparent between the IgH-positive cells and the melanomacrophage cells (compare locations of dark blue cells in A and C to yellow cells in B and D, respectively). In a 15 µm thick section from channel

catfish spleen chromogenically-labeled by ISH for IgH mRNA (E), the denser white pulp (W) can be easily distinguished from the looser red pulp (R) due to background staining [Fange, 1985; Grizzle, 1976; Saunders, 2010]. M = Melanomacrophage center; R = Red pulp; V = Vessel; W = White pulp. These images were obtained from separate regions of the spleen from one catfish. Scale bars represent 50 μm.

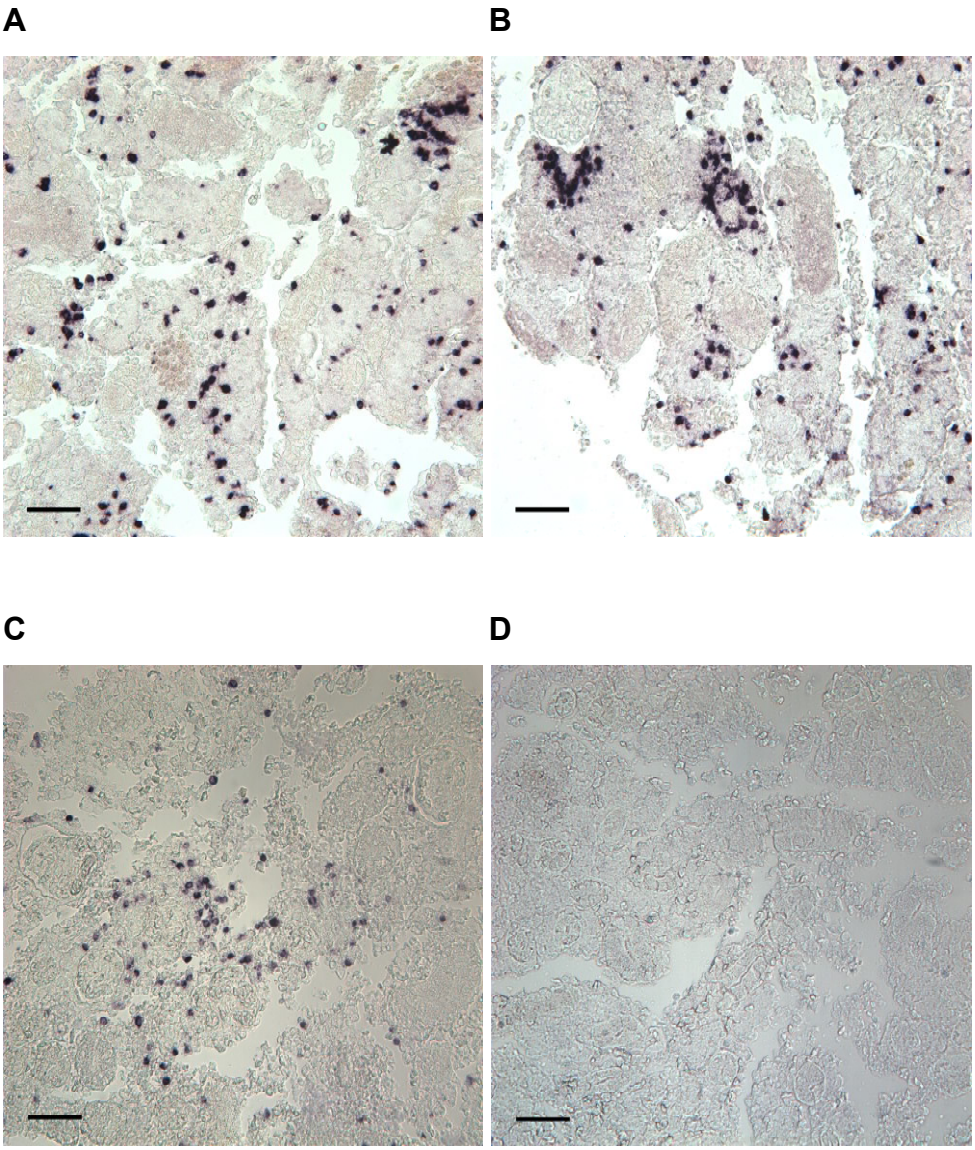


Figure 15. Channel catfish anterior and posterior kidney sections exhibit differences in the distribution of IgH-positive cells.

In channel catfish anterior kidney sections (A and B), dense clusters of IgH-positive cells (dark blue) were a common feature. In channel catfish posterior kidney sections (C), IgH-positive cells were more loosely distributed. Image D is an ISH control in which the sense-strand probe was used in a posterior kidney section. These images were obtained from separate regions of the anterior and posterior kidney from one catfish. Scale bars represent 50 μm.

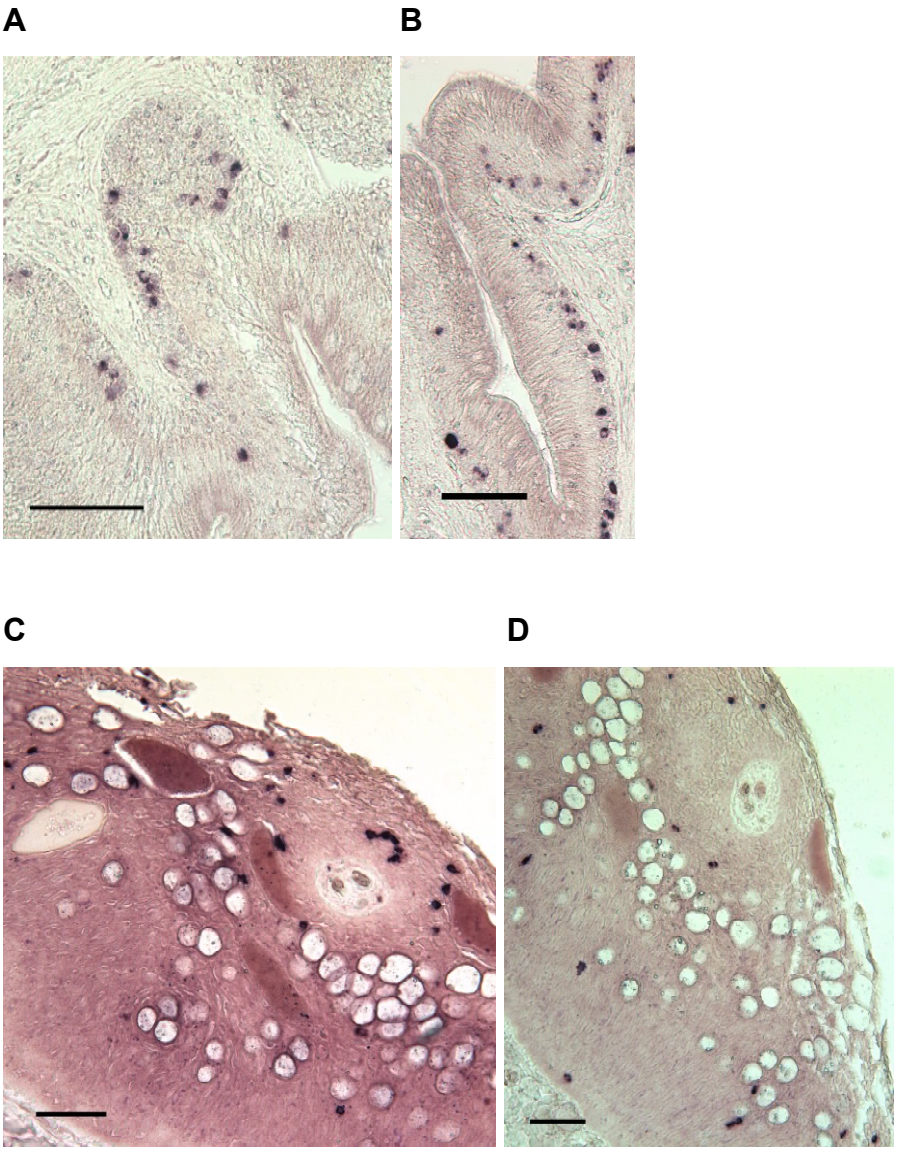


Figure 16. Channel catfish mucosal epithelial tissues exhibit differences in the distribution of IgH-positive cells.

In sections from channel catfish posterior intestine (A and B), IgH-positive cells (dark blue) are located in the mucosa along the basal surface of the lamina propria immediately adjacent to the muscularis mucosae. In channel catfish skin (C and D), which is also a mucosal epithelial tissue in these organisms, IgH-positive cells are distributed throughout the mucosa and tend to occur more toward the apical surface and less often near the basal surface. These images

were obtained from separate regions of the posterior intestine and skin from one catfish. Scale bars represent 50 μm.

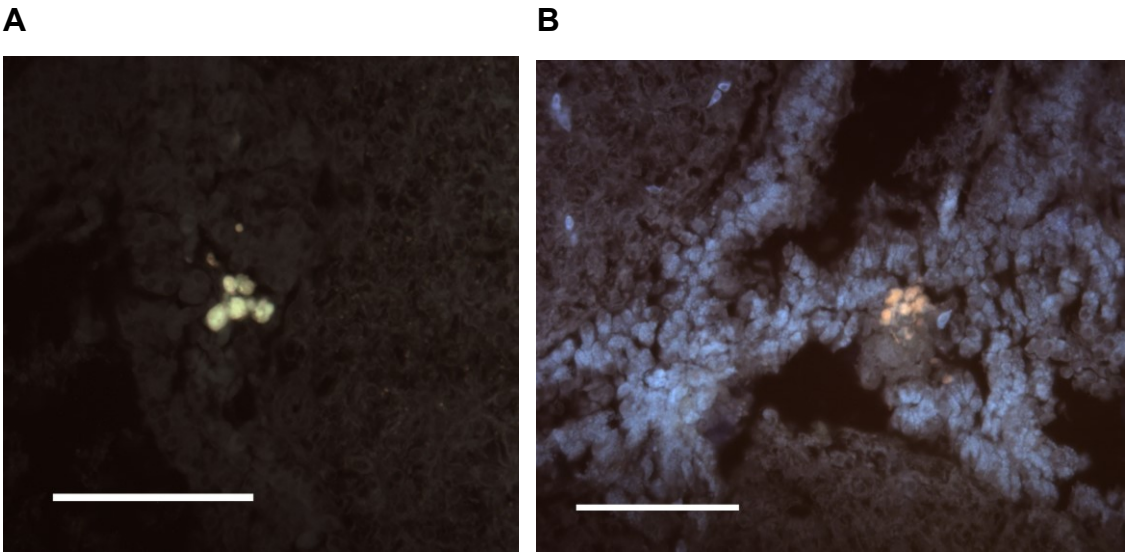


Figure 17. Channel catfish liver contains melanomacrophage cells but no detectable B-cells.

Although the liver is a non-lymphoid tissue, it does have melanomacrophage centers comprised of a few to several cells (autofluoresce yellow under UV channel, A and B). Melanomacrophage cells are recognized to be multifunctional and those of the liver may be engaged in some functions that differ from those found in lymphoid tissues [Agius, 2003; Fange, 1985]. Consistent with RT-PCR data, no IgH-positive or AID-positive cells were identified in channel catfish liver sections by ISH. These images were obtained from separate regions of the liver from one catfish. Scale bars represent 50 μm.

3.3.4 AID-Positive Cells Appear To Preferentially Associate With Melanomacrophage Cells

Although the detection method that culminated in NBT/BCIP staining (i.e. enzyme-linked anti-DIG Fab fragments with chromogen) appeared to be very effective for detection of hybridized IgH riboprobe, it did not provide the degree of sensitivity necessary to distinguish cells in which the AID riboprobe had hybridized. The considerably lower cellular concentration of AID mRNA transcript was likely the main reason that the NBT/BCIP staining method was insufficient for detection of hybridized AID riboprobe, but the background staining and probe-independent chromogenic enhancement of melanomacrophage cells that this detection method generated also contributed to the failure. Indeed, early attempts to identify AID-positive cells with the NBT/BCIP staining method suggested that the cells might be associated with melanomacrophage centers, but probe-independent chromogenic enhancement of these regions even in control sections rendered any conclusions questionable [Saunders, 2010]. Immunostaining of AID proteins could not be used to corroborate these early observations as the ISH for mRNA approach was chosen because there was no anti-channel catfish AID antibody available and it is often difficult to detect AID protein [Poltoratsky, 2007].

To overcome the problems stemming from the low signal and limit chromogenic background staining, a detection system that allowed for signal amplification (i.e. primary anti-DIG mouse antibody and fluorochrome-labeled secondary anti-mouse goat antibody) and utilized a fluorochrome dye (Alexa Fluor 488) for signal detection was used for detection of hybridized AID riboprobe.

In the channel catfish spleen, the AID-positive cells were found in the white pulp and typically occurred in groups that were associated with melanomacrophage cells or centers (Figure 18) [Saunders, 2010]. Although AID-positive cells were found to be associated with melanomacrophage cells, the putative morphological analogues of the mammalian antigen-trapping FDCs, in

contrast to the consistent dark/light zone-polarized, egg-shaped germinal centers which harbour AID-expressing cells in higher vertebrates, the arrangements of AID-positive cells with melanomacrophage cells in channel catfish spleen were more varied. However, even though no standard pattern of colocalization of AID-positive cells and melanomacrophage cells was found in the channel catfish spleen, instances of AID-positive cells surrounded by melanomacrophage cells in an approximately circular organization were more commonly observed than other arrangements, such as AID-positive cells interspersed with melanomacrophage cells, somewhat aligned with melanomacrophage cells but not surrounded by them, or associated with only a few melanomacrophage cells (Figure 18).

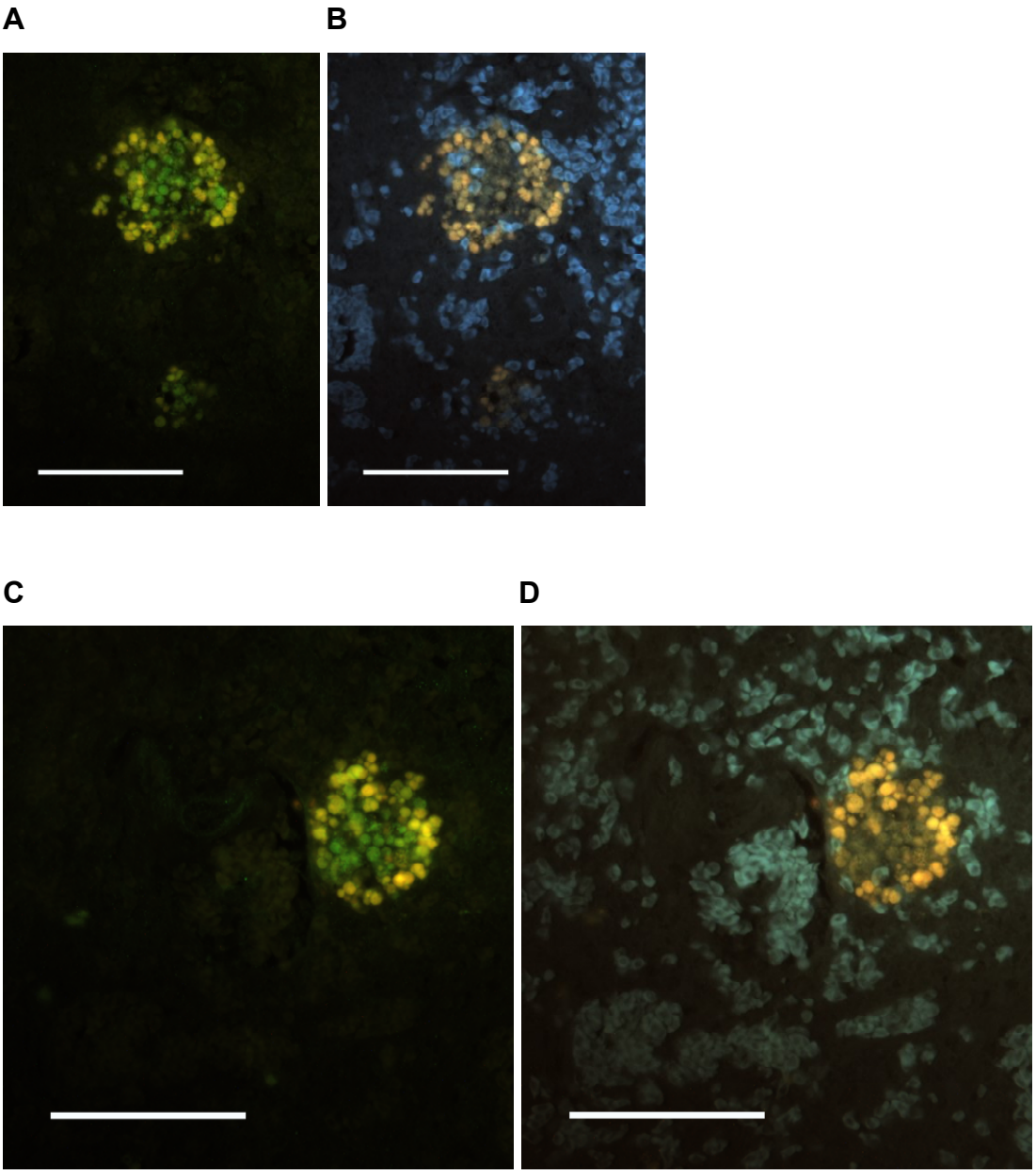
The arrangements of AID-positive cells and melanomacrophage cells were also varied in their overall size, which was attributable not only to the total number of cells present — usually a few hundred to a few thousand cells — but also to their spatial density as some of the arrangements were more compact with cells tightly grouped together while others were more spread out with cells loosely grouped together (Figure 18). In general, it appears that the cell numbers of colocalized AID-positive cells and melanomacrophage cells exhibit a somewhat loose positive correlation; that is, among the associations of AID-positive cells with melanomacrophage cells, if a relatively large number of one of the two cell types is found within a particular association, regardless of its arrangement, then a relatively large number of the other cell type is likely to also be found within the same association. It is not possible to determine from the ISH whether this is influenced by the pre-existing size of the melanomacrophage centers (e.g. larger melanomacrophage centers may be able to support more AID-positive cells than smaller melanomacrophage centers) or other factors (e.g. proportionate recruitment of melanomacrophage cells to accumulations of AID-positive cells).

Occasionally, melanomacrophage centers did not appear to be associated with any AID-positive cells, even in serial sections, despite the fact that these melanomacrophage centers were in proximity to others that did have AID-positive cells associated with them. Likewise, rare occurrences of apparently

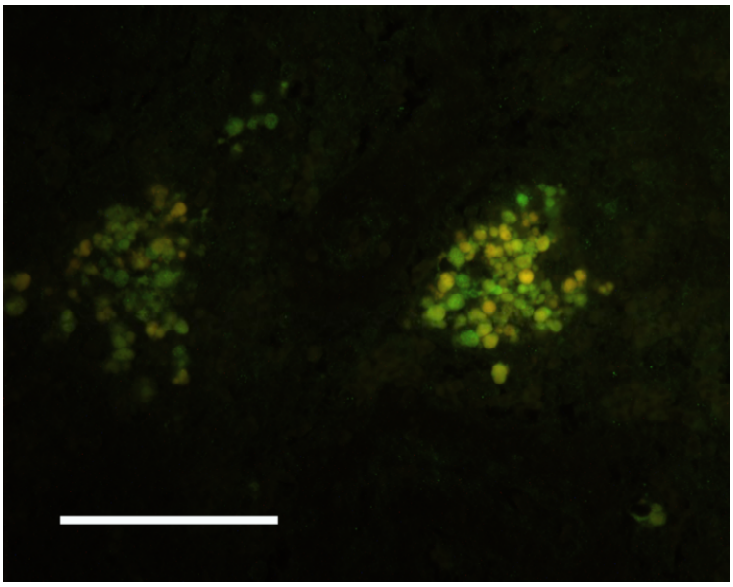
orphaned AID-positive cells were also observed. Lack of a sufficient number of flanking serial sections precluded definite conclusions about whether any of these rare occurrences represent a genuine circumstance of an AID-positive cell(s) without associated melanomacrophage cells or are an artifact of the focal plane (i.e. no z-dimension), but in instances where an adjacent serial section was available, overlying representatives of the other cell type were sometimes found.

AID-positive cells were also identified in the channel catfish kidney (Figure 19) and intestine (see [Saunders, 2010]) in similar associations with melanomacrophage cells as were identified in the spleen. The arrangements of AID-positive cells and melanomacrophage cells in the kidney were much the same as those found in the spleen. In the intestine, however, the associations did not display the same array of arrangements as those found in the spleen and kidney and instead consisted of only a few interspersed representatives of each of the two cell types.

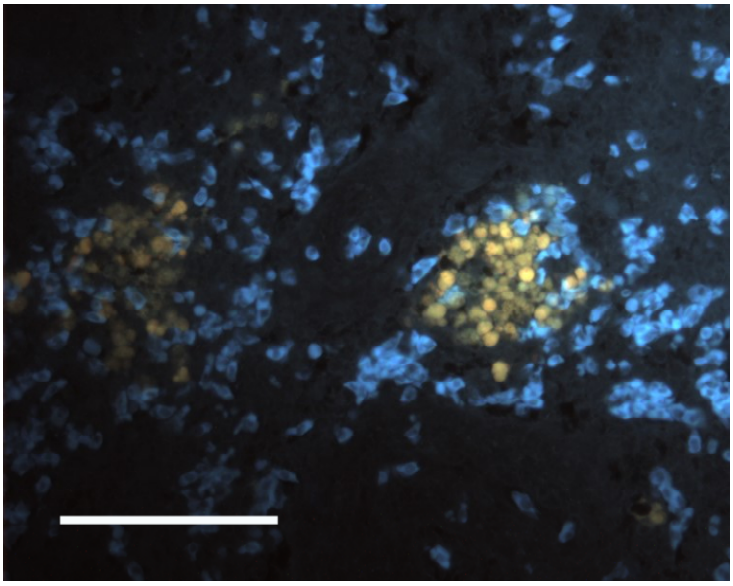
Consistent with RT-PCR data, no AID-expressing cells were identified in the channel catfish liver, even though the liver does harbour melanomacrophages cells (Figure 17).


A
B
C
D

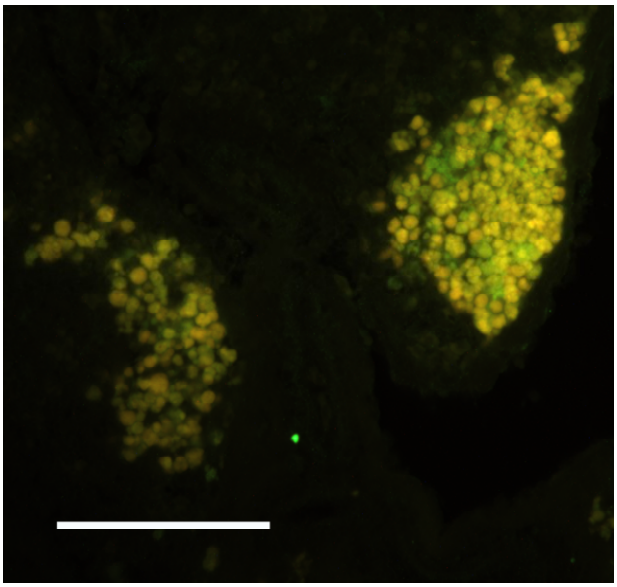
E



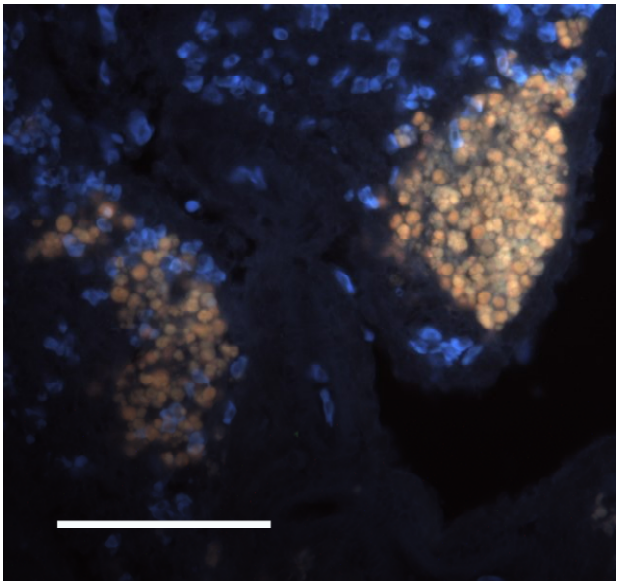
F



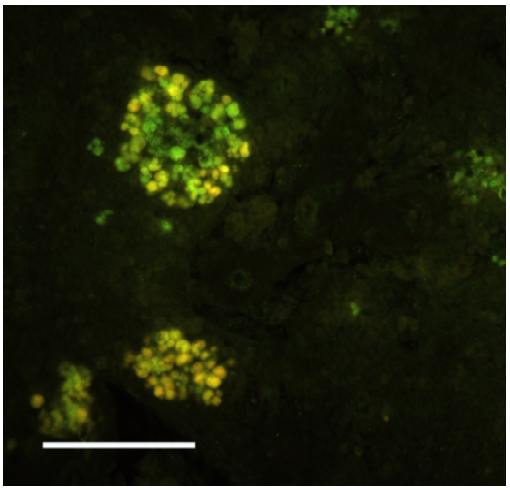
G



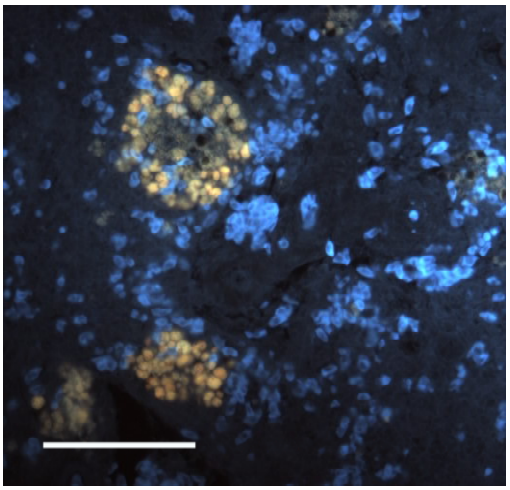
H



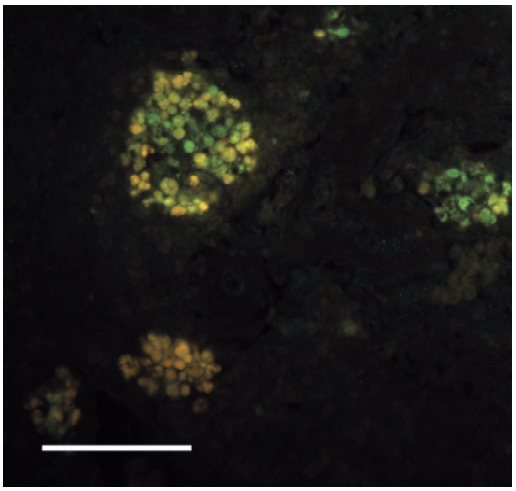
I



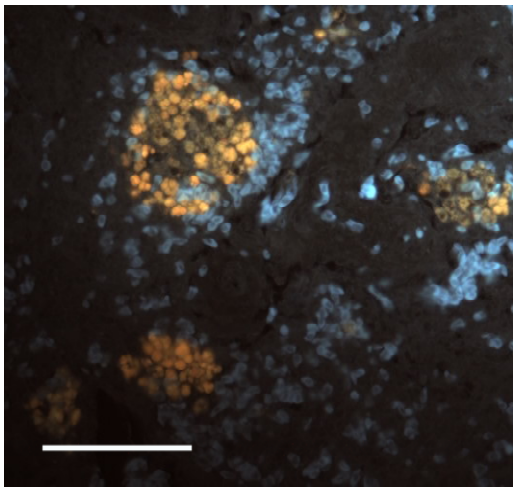
J



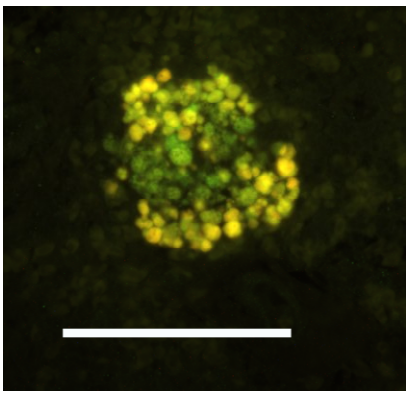
K



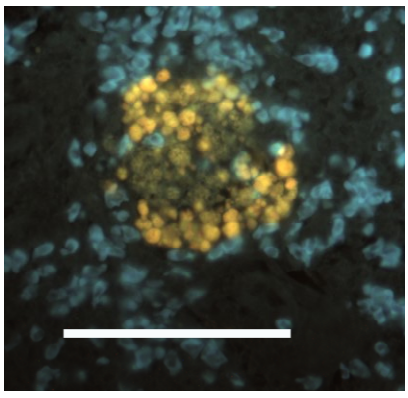
L



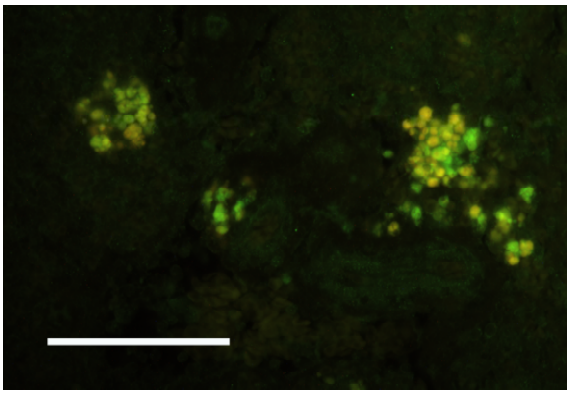
M



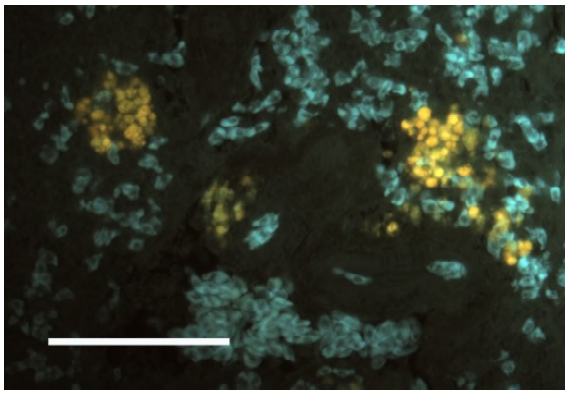
N



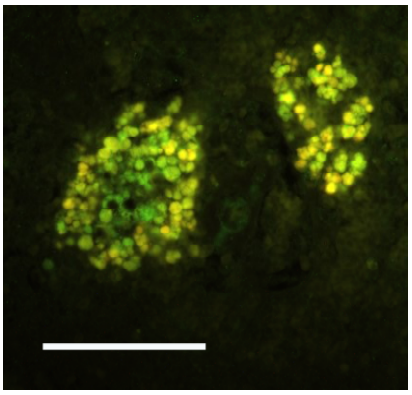
O



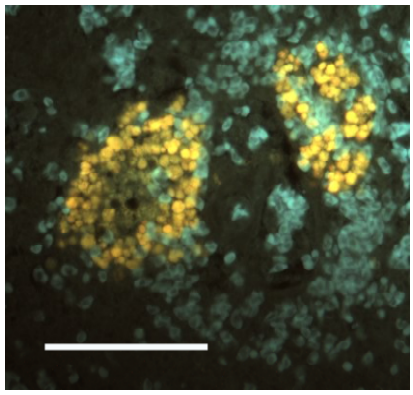
P



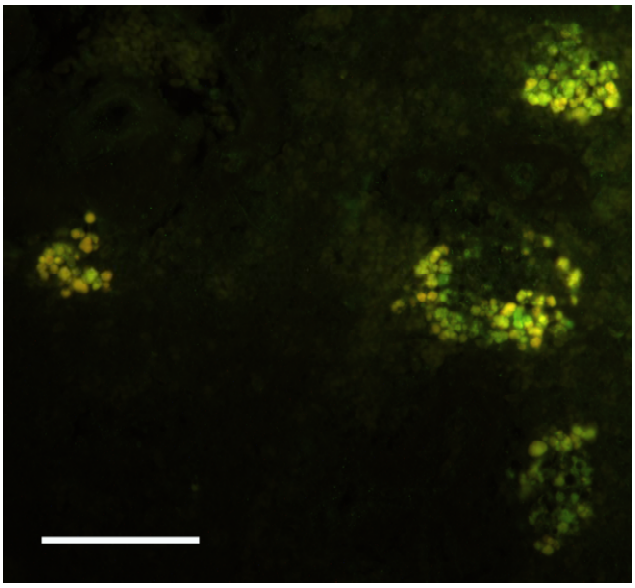
Q



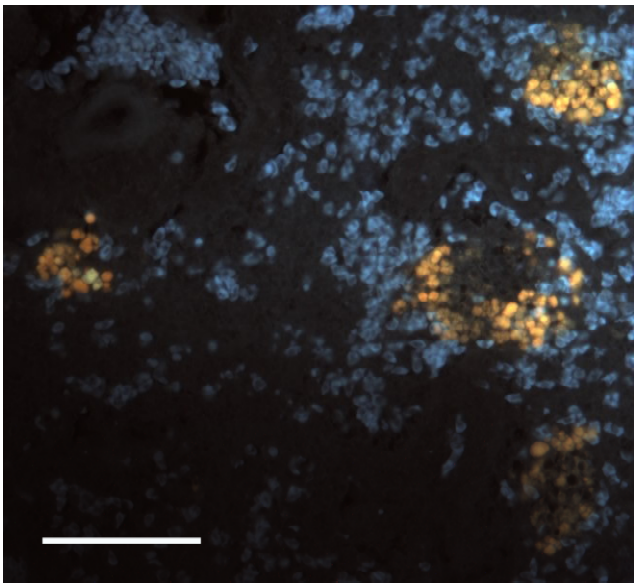
R



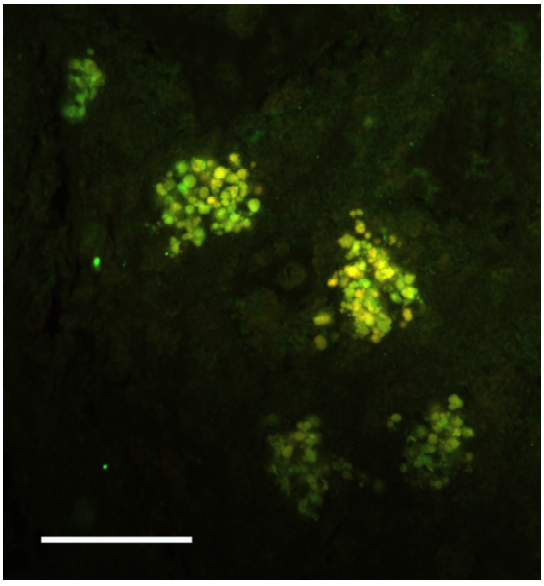
S



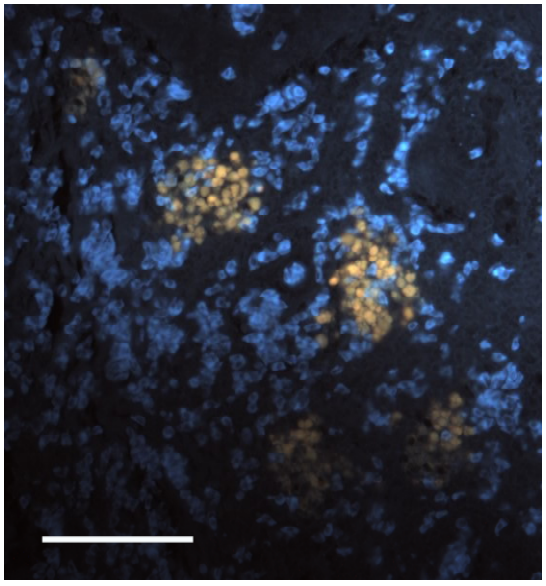
T



U



V



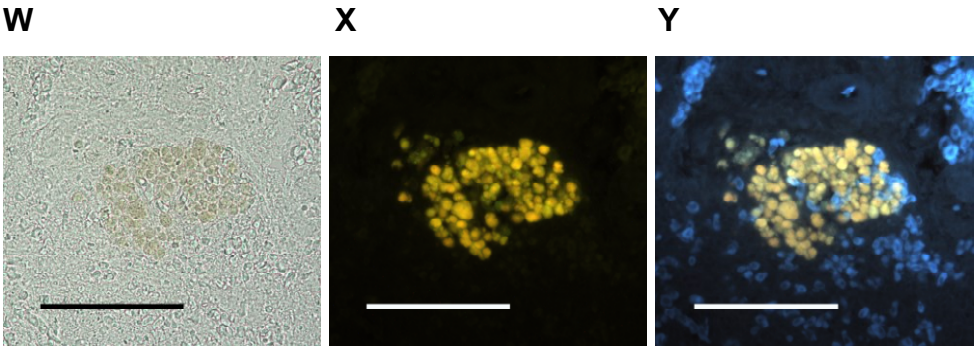


Figure 18. Cells that express AID mRNA appear to preferentially associate with melanomacrophage cells in channel catfish spleen.

In channel catfish spleen sections, ISH for AID mRNA revealed that AID-positive cells (fluorescent green under green fluorescence channel in A, C, E, G, I, K, M, O, Q, S, and U) appear to associate with melanomacrophage cells (yellow under the green fluorescence channel [A, C, E, G, I, K, M, O, Q, S, and U] and UV channel [B, D, F, H, J, L, N, P, R, T, and V]). Although a core of AID-positive cells mostly surrounded by melanomacrophage cells at the periphery in an approximately circular organization (e.g. as seen in C) seems to be a common arrangement, there does not appear to be a standard pattern of colocalization between AID-positive cells and melanomacrophage cells. Images I/J and K/L are of the corresponding region in serial sections. In images of a representative melanomacrophage center (bright field [W], green fluorescence channel [X], and UV channel [Y]) it is clear that no cells were labeled by control ISH using AID sense-strand probe (X). With the exception of the noted serial sections, these images were obtained from separate regions of the spleen from one catfish. Scale bars represent 50 μm.

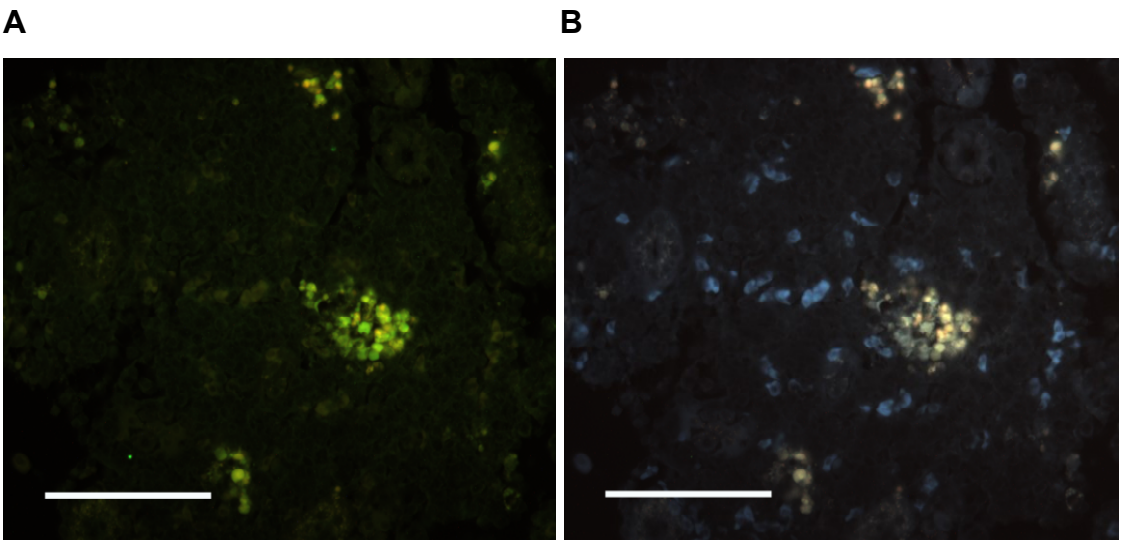


Figure 19. Channel catfish anterior kidney harbours cells that express AID mRNA.

The AID-positive cells (fluorescent green in A) colocalize with melanomacrophage cells (yellowish in A and B) in the channel catfish anterior kidney. Scale bars represent 50 μm.

3.4 Functional Characterization Of AID Expression In A Channel Catfish Immortal B-Cell Line

3.4.1 Mutation Occurred In The IgH Gene Of A Channel Catfish Immortal B-Cell Line After Mitogen-Induced Upregulation Of Endogenous AID mRNA Expression

The initial fragment of the channel catfish AID homologue was RT-PCR amplified from the channel catfish immortal B-cell line 1B10 after it had been stimulated with LPS. Further investigation of 1B10 and another channel catfish immortal B-cell line, 3B11, using semi-quantitative RT-PCR with intron-spanning primers showed that 1B10 does constitutively express AID mRNA at a low level but 3B11 does not constitutively express AID mRNA at a detectable level (Figure 20).

Although the 1B10 cell line had been in continuous culture for more than 10 years at the time of this experiment, the one rearranged IgH VDJ exon had not acquired any mutations. This lack of mutation, despite a relatively prolonged period in culture during which AID mRNA was constitutively expressed at a low-level and the IgH VDJ exon was actively transcribed (Figure 20), is consistent with the conclusion from mammalian studies that AID expression must exceed an undefined threshold before mutation is observed [Zhang, 2001]. Thus, to examine the outcome of AID function in the VDJ exon of the IgH gene in channel catfish immortal B-cells, it was first necessary to determine if the level of AID expression in the cells could be increased to possibly surpass this threshold. To this end, cultures of 1B10 and 3B11 cells were exposed to a variety of substances (i.e. mitogens) known to stimulate channel catfish immune cells to enter an activated state, as AID is expressed by activated B-cells. It is important to note that, while *in vitro* mitogenic activation can result in the upregulation of a variety of genes — based on the particular mitogen and the cell type and state — it is a generalized hyper-activation that is not directly equivalent to the specific physiological activation that occurs *in vivo*.

Various mitogens (LPS, anti-catfish IgM antibody, PMA, and CI) were investigated for their ability to upregulate AID expression in the 1B10 and 3B11 channel catfish immortal B-cell lines (Figure 20). In all instances, 1B10 was found to have more AID mRNA expression than 3B11, which had no detectable AID mRNA expression under any treatment. The combination of PMA/CI was found to be the most potent cell activator for AID mRNA upregulation among those tested. Treatment with ConA, a lectin that is known to act as a mitogen of T-cells but not B-cells, was included as control [Bly, 1986; Lin, 1992].

Interactions with cognate T-cells and the soluble factors they produce are an important component of B-cell activation *in vivo*. While it was not feasible to attempt to replicate the physical interactions with cognate T-cells in this experiment, it was thought that the medium in which a channel catfish immortal T-cell line had been grown might contain soluble factors that could have a synergistic activating effect on the channel catfish immortal B-cells when combined with the mitogens. Therefore, this mitogen experiment was duplicated with the addition of 25% channel catfish immortal T-cell line 28S.3 cell-conditioned medium to attempt to increase expression of IgH (i.e. increase AID access to the IgH gene) but no change in IgH mRNA or BCR were detected via semi-quantitative RT-PCR or FACS, respectively (not shown).

Since teleost fish are known to have slower and much lower antibody affinity maturation than mammals and because the experimental system did not select for mutants (∴ rare mutants may be diluted by non-mutants), the cells were maintained under AID-upregulating stimulation for a prolonged amount of time to ensure sufficient opportunity for mutations to occur at a level that could be readily detected (i.e. sufficient time for expansion of mutant clonal genealogies [Neuberger, 1995]). Exposure to successive rounds of activating agents has been found to be a successful strategy for increasing mutation frequency in cultured mammalian cells [Tarlinton, 1998].

Prolonged culture of repeatedly PMA/CI stimulated 1B10 cells led to the development of mutations in the endogenous IgH gene VDJ exon. Preliminary sequence data revealed 13 mutations (9 are novel; Figure 21) in 14,238 bp

sequenced from 42 subclones of gDNA that had been PCR amplified with Pfu polymerase. One clone had a doublet mutation and the remaining mutations occurred individually in clones. For the purposes of counting mutations herein, the doublet mutation is presumed to be the result of two sequential mutational events in clonal progeny, but it could not be ascertained from the available sequence data (i.e insufficient data to develop lineages, if present) whether this is truly the case or if the doublet mutation is the outcome of a single tandem mutational event.

Even though the small sample size makes statistical analysis impractical, the locations of the mutations are of interest. One of the mutations occurred within the portion of sequence contributed by the D sub-exon (i.e. CDR3) and the others within or at the boundaries of the framework regions FR2 and FR3. Six of the nine novel DNA mutations, which include one of the mutations in the doublet, would result in amino acid replacement in the variable region of the expressed IgH. The mutation in the CDR3 is among the six replacement mutations.

This preliminary data is consistent with channel catfish AID acting as a mutator of the endogenous IgH gene in the context of channel catfish B-cells when expressed above an undetermined threshold and, thereby, establishes that channel catfish are SHM competent. This substantiation of the mutator function of channel catfish AID and the occurrence of SHM in the IgH gene provided justification for the establishment of a controlled AID-expression system that could be used to examine SHM and CSR processes with AID from various organisms in fish B-cells without the need for the addition of mitogens, which non-selectively upregulate many genes and can stress cells (i.e. act as stimulants of apoptosis and as tumor promoters [Lin, 1992]).

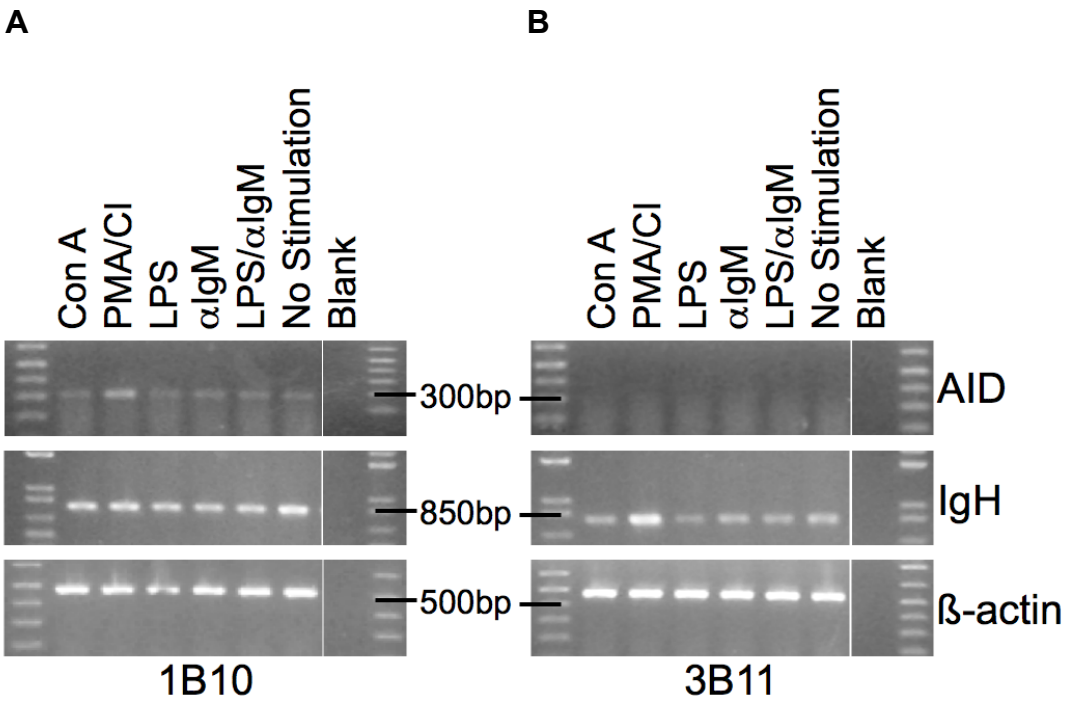


Figure 20. Expression of AID mRNA can be induced in channel catfish immortal B-cell lines.

Semi-quantitative RT-PCR with intron-spanning primers reveals that constitutive expression of AID differs among channel catfish immortal B-cell lines (A and B), as does the extent to which AID expression is upregulated in response to stimulation with different treatments (indicated above gel lanes). The 3B11 channel catfish immortal B-cell line, which expresses both IgM and IgD, is postulated to have arisen from a naïve B-cell. The 1B10 channel catfish immortal B-cell line, which expresses only IgM, is postulated to have originated from a previously activated B-cell. Treatment with ConA, a lectin that is known to act as a mitogen of T-cells but not B-cells, was included as control. Number of PCR cycles: AID 30; IgH 20; ß-actin 25. (ConA: Concanavalin A; PMA/CI: Phorbol Ester/Calcium Ionophore; LPS: Lipopolysaccharide; αIgM: mouse IgG anti-catfish IgM; Blank: no template added to PCR reaction). This experiment was repeated with culture medium that contained 25% channel catfish immortal T-cell line 28S.3 cell-conditioned medium, but no differences in expression were observed. Modified from [Saunders, 2010].

Figure 21. Mutation occurred in a channel catfish B-cell line after upregulation of native AID.

Upregulation of native AID expression in the channel catfish 1B10 immortal B-cell line, via mitogen stimulation (PMA/CI), led to the appearance of mutations in the VDJ variable exon of the endogenous IgH gene. Sequence data revealed 13 mutations (9 are novel) in 14,238 bp sequenced from 42 subclones of gDNA that had been PCR amplified with Pfu polymerase. One clone had a doublet mutation and the remaining mutations occurred individually in clones. A modified version appears in [Dancyger, 2012].

4. Discussion

This research demonstrated that teleosts do possess an AID homologue — a key element of antibody affinity maturation in mammals — and have the capacity for Ig gene mutation in response to sufficient levels of AID expression. Moreover, it revealed that AID-expressing cells occur almost exclusively in association with melanomacrophage cells, a feature which suggests that this organization may have a functional purpose. While this distinctive organization superficially resembles that of the mammalian germinal center, it remains to be determined if it functions similarly (e.g. supports selection).

Firstly, channel catfish express a markedly conserved AID homologue in a tissue expression pattern that is consistent with a role in immune function. That mutation occurred in the variable region of the endogenous IgH gene of a channel catfish immortal B-cell line subsequent to upregulation of native AID is consistent with channel catfish AID being catalytically active (i.e. mutator function) and indicates that their IgH gene organization supports SHM. As noted in the Introduction, SHM of the Ig genes appears to be necessary to achieve substantial increases in antibody affinity via selection in mammals.

Secondly, in channel catfish sections, cells with AID mRNA expression colocalized in discrete clusters with antigen-trapping melanomacrophage cells, which are putative teleost FDC analogues. The consistent co-occurrence of these two cell types suggests that some sort of functional interaction(s) may be occurring between them. That channel catfish are SHM competent and appear to exhibit a tissue architecture that is remarkably reminiscent of mammalian germinal centers suggests that they have the potential for antibody affinity-based selection; however, this is entirely speculation as the actual occurrence of selection has not been demonstrated.

4.1 Channel Catfish Is Capable Of SHM

The identification of an AID homologue in the channel catfish is a milestone in teleost immune research as it is the first time that an AID homologue has been

cloned from any poikilothermic vertebrate [Saunders, 2004]. Its identification in a lower vertebrate model organism confirms the presence of a consistent mechanism for SHM in the antibody genes of jawed vertebrates [Diaz, 1998; Hsu, 2006], which reinforces contemporary theories on the evolution of the gnathostome adaptive immune system (i.e. that the essential components were in place by the advent of cartilaginous fish [Diaz, 1998; Hsu, 2006]). It also provides a definitive means with which to further examine antibody affinity maturation processes in these organisms [Dancyger, 2012; Saunders, 2010; Wakae, 2006].

Subsequent phylogenetic analysis by Conticello et al. [Conticello, 2005] has revealed that the AID protein is one of the earliest members of the AID/APOBEC family of proteins. The evolutionary appearance of AID significantly predates that of APOBEC1, its better-known comparator in mammals, which is absent in teleosts and so precludes AID-APOBEC1 homology comparisons therein [Conticello, 2005; Neuberger, 2003]. Phylogenetic analysis also confirmed that teleost AIDs cluster together and with mammalian AIDs when compared to other eukaryotic deaminases, including the APOBECs [Conticello, 2005]. But homology comparisons of the putative channel catfish AID protein with its mammalian counterpart revealed that, despite being relatively well conserved overall, the putative catalytic active site of the protein is less conserved than the rest of the protein, partly due to the presence of extra amino acids [Saunders, 2004]. These extra amino acids, with some sequence differences, are likewise found in the putative catalytic active sites of the predicted AID proteins of other teleosts, zebrafish (confirmed by [Zhao, 2005]) and Fugu [Conticello, 2005; Saunders, 2004].

Multiple extra amino acids are not found in the putative catalytic active site of the predicted AID protein fragments of Dogfish [Conticello, 2005] and Elephant Shark (Accession # AAVX01329030.1), representative cartilaginous fish, or *Xenopus* (one extra amino acid in catalytic active site) [Conticello, 2005; Ichikawa, 2006; Zhao, 2005], a representative early tetrapod. These organisms are from classes that evolutionarily flank the common ancestor of mammals and

ray-finned fishes [Hsu, 2006; Kumar, 1998] and, like the teleosts, exhibit relatively low levels of antibody affinity maturation [Diaz, 1998; Diaz, 2001; Hsu, 1998]. This suggests that the extra amino acids in the putative catalytic active site of the predicted AID proteins of teleosts are likely not the primary reason for their limited affinity maturation. In support of this surmise, Ichikawa et al. [Ichikawa, 2006] predict that these extra amino acids fall on the “exterior” surface of the fish AID molecule, away from the active site, such that they should not affect its catalytic activity. However, Quinlan et al. [Quinlan, 2017] suggest that they may be located at the dimerization interface and speculate that this “may alter the dimer structure” and thereby impact activity.

It is also apparent from comparisons with the putative AID homologues of cartilaginous fish and early tetrapods that the extra amino acids observed in the catalytic active site of teleost AID arose specifically in the lineage of ray-finned fishes sometime after their divergence [Conticello, 2005]. It is possible that this was permitted by polyploidy, which is common in fishes, or a teleost-specific whole genome duplication event that occurred in the ray-finned fishes branch [Conticello, 2005; Hurley, 2007; Santini, 2009; Volff, 2005]. Genome duplication events can “accelerate” rates of evolution [Hurley, 2007] even though they are not necessarily the primary driver of diversification [Santini, 2009]. The conjecture that polyploidy or genome duplication permitted the evolution of extra amino acids in the catalytic active site of teleost AID, despite apparently strong conservation of this region among other known vertebrate AIDs, is partly based on the premise that whole genome duplication events that occurred earlier in the vertebrate lineage (i.e. two round hypothesis) allowed for the development of conventional gnathostome adaptive immunity [Hsu, 2006]. In fact, immune-related genes, which tend to have a high degree of plasticity [Ichikawa, 2006], are recognized to undergo relatively accelerated evolution and it appears that it is not uncommon for the immune gene homologues of members of the ray-finned fish subclass to be divergent when compared to the rest of the gnathostome vertebrate lineage [Ohta, 2006; Wilson, 1997] and often even when compared amongst themselves [Hsu, 2006]. Compilation of AID homologue sequence data

from several teleost species and other, more primitive, ray-finned fish species will be required to determine when the extra amino acids of the putative catalytic active site appeared in the lineage and to provide better resolution of what might constitute a typical catalytic active site of teleost AID homologues.

Despite the lower degree of conservation in the putative catalytic active site, the channel catfish AID protein appears to maintain most of those amino acids determined from mammalian studies to be essential to SHM [Shinkura, 2004; Ta, 2003], CSR [Durandy, 2006; Durandy, 2005; Minegishi, 2000; Revy, 2000; Ta, 2003], and cytidine deaminase function [Cascalho, 2004; Saunders, 2004], which indicates that it might nonetheless maintain conservation of some of its recognized biological function(s) in the channel catfish. This possibility was further supported by the discovery that different channel catfish immortal B-cell lines are differentially able to upregulate AID expression in response to equivalent mitogenic stimulation (i.e. 1B10 responsive and 3B11 unresponsive to PMA/CI) [Saunders, 2010]. As described in the Materials and Methods, based on their Ig phenotypes, the cell lines examined appear to represent different stages of B-cell maturation — the IgM+/IgD- phenotype of 1B10 is consistent with its having arisen from a previously activated B-cell and the IgM+/IgD+ phenotype of 3B11 is consistent with its having arisen from a naïve mature cell [Wilson, 1997]. Thus, their differential response to equivalent stimuli suggests that AID expression in teleosts, like in mammals, may be specific to both the cell type and stage.

Although the cause of immortalization of the channel catfish B-cell lines is undetermined, the 1B10 B-cell line does not appear to be overtly cancerous (Personal Communication from Norman Miller [U. Miss. Med. Ctr — MMI] to Brad Magor: reintroduction of 1B10 cells to the donor fish did not result in pathology). However, as immortalization is itself indicative of aberration, the characterization of immortalized cell lines is often not a straightforward or definite process and it seems that some do exhibit expression patterns that appear to be inconsistent with their characterized phenotype — of particular note here is the mouse 18-81 cell line which expresses AID and hypermutates but otherwise seems to be a

pre-B-cell line (i.e. a cell stage which should not express AID) [Martin, 2002]. And the fact that different mitogens elicit differential upregulation of AID expression among and within channel catfish immortal B-cell lines (Figure 20) highlights an important caveat of the methodology — activation achieved in this manner is not necessarily equivalent among mitogens nor is it likely to closely resemble *in vivo* physiologic activation. Because ectopic overexpression of mammalian AID in non-B-cells and various organisms has demonstrated that AID expression alone is sufficient to introduce mutations (see Introduction), the failure of mitogenic activation to replicate an intracellular milieu in channel catfish 1B10 B-cells that is equivalent to that of physiological activation in channel catfish B-cells *in vivo* is not an impediment to channel catfish AID studies so long as levels of AID expression are sufficiently upregulated. However, the non-physiologic context should be borne in mind during the interpretation of data as it may impact mutational outcomes.

Despite the shortcomings of mitogenic activation, the results from mitogenic stimulation of channel catfish 1B10 B-cells are in line with an important finding about AID expression levels in higher vertebrates: the level of AID expression must surpass an undefined threshold for AID-mediated outcomes to be observed [Zhang, 2001]. That is, despite constitutive low-level expression of AID, as determined by mRNA analysis, during more than 10 years in continuous culture, the channel catfish immortal B-cell line 1B10 did not acquire any mutations (i.e. exhibit signs of AID action) in the VDJ exon of the expressed IgH gene, yet upregulation of the level of AID expression in these cells, by PMA/CI stimulation, led to detectable SHM in less than three months [Dancyger, 2012]. It is not surprising that no mutations occurred in this B-cell line during its 10 years in continuous culture as the finding is consistent with data from studies with the hypermutating Ramos (human) B-cell line which indicate that prolonged culture periods favour growth of subclones with low or no SHM phenotype, a feature that appears to correlate with the level of AID expression [Zhang, 2001], although other factors have also been found to influence the levels of SHM (e.g. polymerase β expression) [Poltoratsky, 2001].

Phorbol esters and calcium ionophores are known to act synergistically to effect a strong proliferative response in channel catfish B-cells [Clem, 1996; Lin, 1992; van Ginkel, 1994]. That PMA and CI also enhance AID mRNA expression in channel catfish B-cells is not surprising as these mitogens behave as mimics of the physiological second messengers 1,2-diacylglycerol (DAG) and inositol 1,4,5-trisphosphate (IP_3), respectively, which derive from phospholipase C hydrolysis of phosphatidylinositol 4,5-bisphosphate [Blumberg, 1988] during signaling initiated by a variety of receptors including the BCR [Krappmann, 2001; Lin, 1992; Moscat, 2003]. The DAG and IP_3 act synergistically as cofactors to activate various protein kinase C isoforms (PKC) [Blumberg, 1988; Lin, 1992] which then propagate signal cascades that lead, among other things, to the translocation of NF-κB transcription factors from the cytoplasm into the nucleus [Krappmann, 2001; Moscat, 2003] where NF-κB is involved in AID expression [Barreto, 2005; Dedeoglu, 2004]. Both DAG and IP_3 are believed to contribute to activation in channel catfish lymphocytes in a manner equivalent to that found in mammalian lymphocytes [Clem, 1996; Lin, 1992; van Ginkel, 1994], which suggests that, since their biomimics, PMA and CI, upregulated AID in the channel catfish 1B10 B-cells, the signals that initiate AID expression may be similar in these organisms. However, it should be noted that PMA does activate a wider range of PKC isoforms than do DAG and IP_3 [Krappmann, 2001; Moscat, 2003] and is responsible for more diverse responses [Blumberg, 1988].

That SHM occurred in the endogenous IgH gene when mitogenic stimulation with PMA/CI was used to upregulate the level of native AID expression in the channel catfish 1B10 B-cell line indicates that the channel catfish AID homologue likely serves the same SHM function as AID in higher vertebrates, especially when coupled with the channel catfish tissue expression data, which showed that AID is expressed *in vivo* in a pattern that is consistent with its having an immune role [Saunders, 2004]. But, although conservation of function with mammalian AID presents itself as an obvious interpretation of the channel catfish AID data, the requirement for the addition of exogenous chemicals (i.e. PMA and CI) in

order to generate the mutations means that these chemicals must be considered as potential causative agents themselves, especially as they were notably cytotoxic.

The greater stability and potency of phorbol esters and calcium ionophores increases their strength as signal generators over that of more typical physiological stimuli (e.g. antigen and LPS [Krappmann, 2001]) [Blumberg, 1988], which are additionally restricted in their signaling capacity by factors such as receptor availability. In fact, the use of PMA and CI was empirically determined to be a necessary choice to achieve the requisite increase in expression of AID mRNA in the 1B10 B-cells (Figure 20) [Saunders, 2010]. However, despite the fact that their mitogenic activity is so potent that they are considered to be tumour promoters, especially the phorbol esters [Blumberg, 1988], the mitogens themselves were not likely the source of the mutations observed in the IgH gene as they are not believed to be mutagenic themselves [Blumberg, 1988]. And subsequent *in vitro* [Dancyger, 2012; Wakae, 2006] and *in vivo* [Yang, 2006] data (discussed below) provided further support for the initiation of hypermutation by channel catfish AID and of genuine SHM in channel catfish.

Some PCR-amplified IgH product from the 1B10 PMA/CI stimulation experiment was provided to later collaborators. They wanted more unselected mutation data from the 1B10 B-cells to complement cell-free *in vitro* reaction data that indicated channel catfish AID doesn't utilize some WRC motifs with the same relative activity levels as human AID or zebrafish AID, even at channel catfish physiological temperatures [Dancyger, 2012]. They extended the section sequenced to include the region 5' of VDJ CDS to the transcriptional start site (see Supplemental Figure S2 B of reference [Dancyger, 2012]). This effort revealed a further six novel mutations, beyond the nine novel mutations identified from the previous sequencing of the 1B10 VDJ CDS (see Results).

Although the examination of a single VDJ sequence has the potential to significantly bias mutation data [Poltoratsky, 2001] and the paucity of mutation data obtained from the IgH VDJ of the PMA/CI stimulated 1B10 B-cells

[Dancyger, 2012] precludes all but superficial qualitative comparisons with the *in vivo* channel catfish IgH VDJ SHM data [Yang, 2006], it appears that the characteristics of the mutations obtained from the IgH gene of the channel catfish 1B10 B-cell line [Dancyger, 2012] were generally consistent with those obtained from channel catfish IgH genes *in vivo* by Yang et al. [Yang, 2006]. Similar to mammalian SHM, the SHM mutation hierarchy among bases in channel catfish is mutation at G≈C>A>T, there is a predilection for transitions over transversions, preferential targeting to RGYW/WRCY hotspots occurs, and mutations occur throughout the entire VDJ sequence of the IgH gene [Dancyger, 2012; Wakae, 2006; Yang, 2006].

There were two apparent disparities noted in the findings from the channel catfish. First, although channel catfish AID targets the RGYW/WRCY hotspot motif like mammalian AIDs [Wakae, 2006; Yang, 2006], it appears that the hotspot motif preferences of channel catfish SHM may be represented by a slightly more “restricted” set of WRC motifs than those of human SHM [Dancyger, 2012; Yang, 2006]. That is, certain dinucleotide or trinucleotide sub-motifs do not appear to be mutated as frequently in channel catfish as they are in humans. Because several processes contribute to SHM outcomes (e.g. repair and selection), more research will be needed to determine the reason for this apparent minor variation in sub-motif preference. It is possible that it is merely an artefact of analyzing a relatively small number of mutations from a relatively small number of gene sequences. However, that targeting by AID might contribute to it is supported by an *in vitro* experiment (i.e. system without repair or selection) that examined the action of AID molecules from various species on some “bubble” substrates and found that not all were utilized as relatively well by channel catfish AID as they were by human AID [Dancyger, 2012]. That channel catfish AID might exhibit minor variation in finite targeting specificity is not inconceivable as some mutations in human AID have been found to affect its “deamination specificity” [Bransteitter, 2004; Bransteitter, 2006] and, despite being markedly conserved, the channel catfish AID sequence does differ from the human AID sequence. Indeed, Diaz et al. [Diaz, 2001] suggest that, given the overarching

similarities and minor differences in SHM from sharks to mammals, evolution may have favoured the optimization of different aspects of the process in different lineages while maintaining the core mutational event (i.e. AID-mediated deamination) and certain associated mechanisms [Ichikawa, 2006].

Second, Yang et al. [Yang, 2006] found that the WA motif did not appear to be significantly targeted for SHM in channel catfish IgH genes. Although this initially suggested that hypermutation in teleosts may not have evolved to incorporate the second phase of mutation to the extent seen in higher vertebrates (see Introduction), this was clearly refuted by a later SHM study in zebrafish IgL genes that found that the WA motif is targeted for SHM [Marianes, 2011]. While the zebrafish IgL gene SHM study indicates that phase two mutation does occur in teleosts, the amount of mutation data available from either fish is extremely limited so further research will be needed to determine if there is a genuine difference in the utilization of phase two mutation between the channel catfish and the zebrafish and/or in its recruitment to the IgH gene versus the IgL gene. Differential usage of repair contributes to gene conversion versus SHM in some mammalian species and Ig genes in the nurse shark exhibit different mutational spectra with IgNAR being similar to conventional SHM while mutation in IgM is predominantly GC-focused [Diaz, 2001].

Although antigen-based selection mechanisms are paramount to improving antibody affinity, it is likely that some reduction of the risks associated with mutation is actually accomplished by mechanisms intrinsic to the Ig genes (e.g. hotspots [Barreto, 2005]). Like mammals, the Ig genes of channel catfish and other teleosts have undergone codon optimization [Conticello, 2005; Yang, 2006]. Although SHM occurs throughout the IgH VDJ sequence [Dancyger, 2012], codon optimization favours replacement mutations in the CDR and silent mutations in the FR [Yang, 2006]. This strategy might not only minimize risk but also enhance the likelihood of favourable change [Diaz, 1998]. That there was enough pressure to evolve this strategy suggests that there may also have been pressure to evolve other mechanisms (e.g. selection) to enhance the R/S ratio in CDR versus FR. Despite the clear occurrence of SHM *in vivo* and the

identification of a few very small apparent clonal hierarchies, Yang et al. [Yang, 2006] did not find compelling IgH gene sequence evidence in support of the occurrence of selection in channel catfish — at least not strong conventional selection based on higher R/S ratios in CDR versus FR. Although apparently consistent with genetic data from *Xenopus* [Diaz, 1998; Hsu, 1998] and teleost antibody affinity data (e.g. apparently meager affinity increase) [Cain, 2002; Kaattari, 2002; Meffre, 2001; Yang, 2006], this finding was also contradicted by Marianes and Zimmerman [Marianes, 2011] who identified substantial evidence for clonal lineages among hypermutated zebrafish IgL genes. It is likely that the channel catfish study was compromised by bias introduced in the experimental approach in which a very few representative cDNA sequences (104 total) from each of the thirteen known channel catfish VH families were analyzed [Yang, 2006] as this does not provide a sufficient sequence data set to either definitively exclude the presence of or sufficiently identify indicators of selection (e.g. clonal hierarchies and altered R/S ratios), especially in instances where selection may be weak.

Affinity studies on mammalian antibodies have indicated that, in most instances, significant increases in antibody affinity are achieved through the cumulative effects of several minor alterations of the protein sequence that occur over several rounds of mutation [Wedemayer, 1997] — a finding also supported by mutational enhancement of other non-antibody protein molecules in selective systems [Wang, 2004]. And the benefits of mutation should outweigh the risks only if there is a means to protect against detrimental mutants and, ideally, promote desirable mutants. That channel catfish are able to introduce replacement mutations into the variable region of their IgH genes and teleost antibody affinity does increase in response to repeated antigen exposures suggests that they employ some sort of selective mechanisms to accomplish this.

While the apparently lower frequency of SHM in channel catfish is likely partly due to the relatively lower activity of the channel catfish AID (discussed below) [Dancyger, 2012; Wakae, 2006], it may also reflect a weaker ability to select mutants. Wang et al. [Wang, 2004] suggest that whole genome hypermutation in

AID-expressing B-cells in mammals is more commonplace than generally acknowledged — partly due to the focus on the analysis of specific genes (i.e. lack of sequencing in other genes) — but that strong selection mitigates detrimental mutations even outside of the Ig genes. This hints that selection capabilities may act to cap the catalytic rate of AID during evolution. Indeed, Ichikawa et al. [Ichikawa, 2006] have suggested that evolutionary pressure has selected for “catalytic suboptimality” in AID. Given their apparently weaker ability to select B-cells bearing BCR of higher affinity, the need to balance SHM mutation rate with selection capabilities is a plausible explanation for the lower catalytic rate of channel catfish AID. But the differences between mammals and fish are not limited to selection and so other factors likely also contribute to lower activity in channel catfish AID.

In various experimental systems (e.g. a mammalian cell line with an SHM reporter, bacterial and yeast mutation assays, and a cell-free “bubble” substrate system), it appears that the AID enzymes from different organisms, including teleosts (channel catfish, zebrafish, and Fugu), exhibit different catalytic rates and that these rates are affected by temperature [Barreto, 2005; Conticello, 2005; Dancyger, 2012; Ichikawa, 2006; Wakae, 2006]. Not surprisingly, the various AID molecules tend to have their maximal activity rate within the physiological temperature range of the organism from which they derive [Barreto, 2005; Conticello, 2005; Dancyger, 2012; Ichikawa, 2006; Wakae, 2006]. That is, among the various tested temperatures (intermittent from 15°C to 37°C), the mouse and human AIDs were optimally active at 37°C while the fish AIDs were optimally active at 25°C [Barreto, 2005; Dancyger, 2012; Ichikawa, 2006; Wakae, 2006].

Although there was some variation in the relative activities of the AID molecules depending on the experimental system used, the overall result was principally the same with the maximal activity of the mammalian AIDs exceeding the maximal activity of the fish AIDs [Barreto, 2005; Dancyger, 2012; Ichikawa, 2006; Wakae, 2006]. The sole exception was observed in an *in vitro* cell-free

experimental system wherein the maximal activity of the zebrafish AID was significantly higher than the maximal activity of the human AID on some of the "bubble" substrates (i.e. sequence-dependent) [Dancyger, 2012]. That the relative activity of mammalian AID was typically greater than that of fish AID was not surprising given their differences in antibody affinity maturation capabilities. However, that there is a considerable difference in the relative activities of the various fish AIDs, with the relative activity of zebrafish AID being consistently considerably higher than channel catfish or Fugu AIDs, was unexpected [Barreto, 2005; Dancyger, 2012; Ichikawa, 2006; Wakae, 2006].

Intriguingly, in the *in vitro* cell-free experimental system in which maximal zebrafish AID activity exceeded that of human AID, the activity of zebrafish AID at 37°C approximated that of human AID [Dancyger, 2012]. That AID from at least some teleosts can be as catalytically active as that from a mammal suggests that the diminished antibody affinity maturation of lower vertebrates is not entirely the result of catalytic deficiencies related to the extra amino acids present in the putative catalytic active site. Moreover, exchanging the catalytic active site of human AID for that of channel catfish AID yields a chimeric molecule with an activity that approaches that of wild-type human AID at 37°C [Dancyger, 2012]. This further suggests that temperature optimization was not the main evolutionary pressure for the addition and maintenance of the extra amino acids in the teleost AID catalytic active site. Indeed, there are no extra amino acids in the catalytic active site of AID from representative cartilaginous fish and amphibians (one extra amino acid in *Xenopus*), yet they live within similar temperature ranges to teleosts and display SHM.

It is well recognized that the catalytic capabilities of proteins are rarely stand-alone properties with their activity often modulated by other domains of the protein [Bransteitter, 2006]. That the catalytic active site of channel catfish AID is nearly as active as that of human AID, when couched within the human amino- and carboxy-terminal domains, suggests that features other than the conspicuous differences in the channel catfish catalytic active site must be

primarily responsible for the difference its catalytic activity. Domain exchange experiments revealed that the carboxy-terminus of AID exerts influence over the catalytic activity of the molecule and, although the data is sparse, there is also some indication that a portion of the carboxy-terminus of channel catfish AID may modulate the mutational spectrum (i.e. JP8Bdel leads to a shift from approximately equal mutation at C and G to mutation mainly at G [75%]) [Dancyger, 2012; Ichikawa, 2006; Wakae, 2006]. In the *in vitro* cell-free experimental system, at 37°C, the carboxy-terminus of channel catfish AID damped the activity of human AID while the carboxy-terminus of human AID improved the activity of channel catfish AID [Dancyger, 2012]. Thus it appears that, rather than the conspicuous differences in the catalytic active site (i.e. extra amino acids), the activity of channel catfish AID is damped by seemingly lesser features that reside in the carboxy-terminal domain [Dancyger, 2012].

This effect was also observed with another teleost AID in a bacterial mutation assay. A significant increase in rifampicin resistance was observed when the entire carboxy-terminus of Fugu AID was replaced with that of human AID and a lesser increase in rifampicin resistance was observed when only the NES was exchanged [Ichikawa, 2006]. The reciprocal chimera — human AID with Fugu carboxy-terminus — produced rifampicin resistance similar (slightly less) to Fugu AID [Ichikawa, 2006].

Of particular note in the channel catfish AID carboxy-terminus is the aspartic acid at position 176, which is located adjacent to the putative ssDNA-binding groove of the molecule [Dancyger, 2012]. There is a glycine at the corresponding site in human AID (position 164) and reciprocal mutation of this position leads to a significant decrease in function of human AID and increase in function of channel catfish AID at 37°C [Dancyger, 2012]. Indeed, activity of channel catfish AID D176G improves over wild-type at all measured temperatures (18°C, 25°C, and 37°C) yet maintains the pattern of thermosensitivity seen in wild-type channel catfish AID (i.e. higher activity at physiological temperatures and lower activity at 37°C) [Dancyger, 2012]. It appears that the higher activity of channel

catfish AID D176G is associated with improved substrate binding affinity [Dancyger, 2012].

In light of this and given their relative activity levels (i.e. higher in zebrafish and lower in Fugu), it is perhaps not surprising that at this position zebrafish AID has a glycine, the same as human AID, and Fugu AID has a glutamic acid, an acidic residue similar to channel catfish AID [Dancyger, 2012; Saunders, 2004]. But, when compared to mouse AID, zebrafish AID is slightly better conserved than channel catfish AID, so features other than this glycine likely contribute to the considerable difference in their relative activities [Wakae, 2006]. Intriguingly, this glycine appears to be well conserved except among the teleost AIDs [Dancyger, 2012].

And diminished conservation among teleost AIDs of amino acids that appear to be well conserved in AIDs from all other organisms studied, including a representative cartilaginous fish, is a phenomenon that also impacts other amino acids including the residue position that corresponds to human AID serine-38 and its associated PKA phosphorylation consensus sequence (i.e. RRXS) [Bransteitter, 2006; Conticello, 2005; Dancyger, 2012; McBride, 2006; Saunders, 2004]. In mouse/human AID, phosphorylation of serine-38 contributes to post-translational regulation of AID activity (e.g. targeting) [Mu, 2018]. Alterations at serine-38 and its associated PKA phosphorylation consensus sequence in mammalian AID diminish processivity, which impacts mutation rates, and affect “deamination specificity” [Bransteitter, 2004; Bransteitter, 2006]. At the position that corresponds to mouse/human AID serine-38, both zebrafish and Fugu AIDs have a glycine residue [Conticello, 2005; Dancyger, 2012]. Channel catfish AID actually does have a serine at this position, but the associated PKA phosphorylation consensus sequence is disrupted (i.e. KRNS) [Bransteitter, 2006; Dancyger, 2012; McBride, 2006; Saunders, 2004]. However, in zebrafish AID, an aspartate residue (D44) has been identified as an apparent phosphoserine mimic site (i.e. a negative charge that may mimic phosphorylation) [Barreto, 2011]. While this residue is also present at the equivalent position in channel catfish and Fugu AIDs, the actual impact of this

difference has yet to be assessed in fish AID. In addition to serine-38, several other potential phosphorylation sites have also been proposed to contribute to phosphoregulation of mouse/human AID and analyses of their effects are ongoing.

Overall, it appears that there is a considerable range of activity among the few teleost AIDs analyzed and the activity of teleost AID is typically much lower than the activity of mouse or human AID, even in experimental systems that exclude confounding effects such as post-translational modification and protein-protein interactions. And it seems that the catalytic active sites of the teleost AIDs have the potential to be as active as those of mouse and human AIDs, yet their activity appears to be purposely repressed by other elements of the protein. This supports the concept, mentioned above, that there may be evolutionary pressure to cap the activity of AID [Ichikawa, 2006]. However, whether this evolutionary pressure is exerted as a consequence of selection capabilities during antibody affinity maturation has not been determined.

And teleost AIDs have another intriguing idiosyncrasy. As noted in the introduction, studies of AID from patients with HIGM2 syndrome and from artificially generated mutant AID proteins have identified a region in the carboxy-terminal domain that is crucial for CSR in mammals but dispensable for SHM [Ta, 2003]. The ultimate eleven carboxy-terminal residues contain the NES, which appears to be functional for export as teleost AID exhibits the same cellular sequestration pattern as mammalian AID in both mammalian and channel catfish cells [Ichikawa, 2006; Methot, 2015; Wakae, 2006]. The high degree of conservation in this region in the predicted channel catfish AID protein, partly due to the presence of the NES [McBride, 2004], suggested that it may be capable of CSR function, despite CSR in channel catfish being precluded by the absence of the requisite genetic components (see Introduction) [Saunders, 2004; Wakae, 2006].

Fascinatingly, channel catfish AID is CSR competent in mouse AID-/- B-cells, as are AID homologues from other representative teleosts (i.e. zebrafish and Fugu) [Wakae, 2006][Barreto, 2005]. The extent of the relative CSR activity

varies considerably among the AIDs from the few tested fish with the relative CSR activity of zebrafish AID, which is not quite as active as mouse AID, being considerably higher than that of the channel catfish or Fugu AIDs [Barreto, 2005; Ichikawa, 2006; Wakae, 2006]. Thus, the relative CSR activity of the teleost AIDs appears to be generally consistent with their relative SHM activity [Barreto, 2005; Wakae, 2006].

An examination of chimeric AID molecules revealed that, relative to wild-type channel catfish AID, the CSR generated in mouse AID-/- B-cells is slightly improved (about 2-fold) by the exchange of the carboxy-terminal region of channel catfish AID affected by JP8Bdel (described in Appendix) for its human AID equivalent [Wakae, 2006].

To examine SHM and CSR evolution (i.e. using reporters) in the context of channel catfish cells, ectopic expression of full-length and carboxy-truncated AID from representative teleosts (channel catfish and zebrafish) and a mammal (human) using a Tet-Off Expression System was attempted in a channel catfish B-cell line (see Appendix). But this was not completed due to problems with the expression system that were not resolved due to time constraints.

Intriguingly, despite the absence of CSR in teleost fishes, it appears that sharks, which possess a multicluster IgH gene organization that contains “100-200 miniloci”, can undergo a form of isotype switching that relocates the VDJ exon from one locus to another locus [Hsu, 2016; Magor, 2015]. There is no indication of the chimeric VDJ exon products that would be expected to result from a RAG-mediated switch mechanism (i.e. VDJ recombination between sub-exons in different loci) [Hsu, 2016]. Although canonical switch region sequences are absent in shark IgH gene loci, isotype switch events appear to correlate with AID expression and the occurrence of SHM [Hsu, 2016].

In addition to an alternative form of isotype switching, studies in sharks have revealed that antigen receptor genes other than Ig may be normal physiological targets of diversification by AID-mediated SHM [Magor, 2015; Ott, 2018]. Although SHM does not appear to contribute to the diversification of TCR in humans and mice, it does appear to play a role in generating the primary T-cell

repertoire in some other organisms, including sharks and camels [Ott, 2018]. A study of nurse shark TCRα indicates that AID-mediate SHM of TCR occurs during the T-cell proliferative phase (i.e. after RAG-mediated TCR rearrangement) in the thymus, where T-cells are then subjected to selection [Ott, 2018].

That AID appears to contribute to unorthodox activities in shark immunity, including non-canonical locus switching, provides further support for the hypothesis that teleost AID, which has been demonstrated to be CSR competent, may be able to support CSR in reporter constructs in teleost cells.

4.2 AID-Positive Cells Are Associated With Melanomacrophage Cells

Not only did the mitogenic stimulation of 1B10 and 3B11 cell lines suggest that the cell type and state restrictions of AID seen in mammals may be operational in fish, the tissue mRNA expression pattern of the channel catfish AID — anterior and posterior kidney, spleen, intestine, and skin — is consistent with it's having a role in immune function in the fish [Saunders, 2004]. Additionally, while expression of IgH mRNA was seen in all channel catfish tissues that had expression of AID mRNA, expression of AID mRNA was not seen in all tissues that had expression of IgH mRNA (all tissues examined) [Saunders, 2004] — i*n vivo* confirmation that AID mRNA is expressed only in a subset of B-cells in channel catfish.

The tissue expression pattern of AID appears to be more diverse in channel catfish than it is in mice [Muramatsu, 1999] and humans [Muto, 2000] as the expression of channel catfish AID mRNA was seen not only in secondary lymphoid tissues (spleen, intestine, and posterior kidney) but also in primary lymphoid tissue (anterior kidney) and fin (i.e. essentially skin) [Saunders, 2004]. It has been demonstrated that AID transcripts can be amplified from a variety of non-lymphoid human tissues if a large enough number of PCR cycles are used [Muto, 2000], but this is not likely the cause of the observed diversity in tissue expression of AID in the channel catfish as the PCR cycle numbers were kept

relatively low. While the comparatively wider range of tissue expression of AID in channel catfish might seem inconsistent with the expression profile one might expect based upon the immune paradigm of higher vertebrates, it can be reconciled on closer consideration of each case in the fish.

In both mammals and fish [Tort Bardolet, 2003], certain mucosal tissues (e.g. gastrointestinal tract) are known to contain secondary lymphoid tissue. Given that the skin of channel catfish is a mucosal tissue [Tort Bardolet, 2003] that has intimate contact with the external environment, it is logical to postulate that the relatively high level of AID mRNA expression observed within the channel catfish skin may originate from the presence of secondary lymphoid elements therein [Iwama, 1996; Saunders, 2004].

The interpretation of approximately equivalent levels of AID mRNA expression in both the anterior and posterior portions of the channel catfish kidney is complicated by the dual role of kidney as both a primary and secondary lymphoid organ in fish [Zwollo, 2005] and the unknown potential for AID to contribute to both primary and secondary repertoire diversification in fish as it does in some higher vertebrates [Arakawa, 2002; Diaz, 2002]. Indeed, amongst higher vertebrates there are some differences in AID usage (e.g. initiates SHM versus gene conversion) and in some it contributes to antigen-independent primary repertoire diversification through SHM (sheep and rabbit) and/or gene conversion (rabbit, chicken, swine, and cattle) [Arakawa, 2002; Dedeoglu, 2004; Diaz, 2002; Diaz, 1998; Green, 1998; Neuberger, 1995; Papavasiliou, 2002] in addition to antigen-dependent secondary repertoire diversification [Hsu, 2006]. But this may be a quirk of higher vertebrates. In sharks, SHM does not appear to contribute to primary repertoire diversification [Hsu, 2006]. And research from zebrafish indicates that SHM likely does not play a role in primary repertoire diversification in teleosts [Marianes, 2011].

The partitioning of primary and secondary lymphoid function has not been examined in the kidney of channel catfish, but it has been examined in the kidney of trout where changeover between the two functions was found to occur along an anterior-posterior gradient with primary lymphoid function (i.e. B-cell

development) mainly in the anterior portion of the kidney and secondary lymphoid function (i.e. B-cell maturation) mainly in the posterior portion of the kidney [Zwollo, 2005]. But the most recent common ancestor of channel catfish and trout dates to approximately 180 million years ago [Santini, 2009] and ray-finned fish are quite functionally diverse [Saunders, 2004] so nominal differences in the distribution of lymphoid function within their respective kidneys are not improbable given their phylogenetic distance. Since the approximately equivalent levels of AID mRNA expression in the anterior and posterior portions of the channel catfish kidney are likely only attributable to secondary repertoire diversification, it may be that the distribution of secondary lymphoid function is more homogeneous in the kidney of channel catfish than in that of trout.

Unfortunately, attempts to resolve this issue by using ISH to explore the spatial relationship of AID-positive and RAG1-positive cells, as markers for B-cells undergoing AID-mediated mutation and developing B-cells (i.e. primary repertoire), respectively, in the anterior kidney were not successful due failure of the RAG1 ISH. This was presumably a consequence of insufficient sensitivity of the detection method employed (corroboration for this supposition comes from RT-PCR for TCR, another target of failed ISH probes, from channel catfish spleen samples obtained via laser-capture microdissection, discussed below). Although RAG1 ISH was unsuccessful, comparison of results from IgH and AID ISH supports that a role for AID in primary repertoire diversification is doubtful in channel catfish and that some amount of secondary lymphoid function takes place in the anterior kidney. That is, there are large numbers of often densely-grouped IgH-positive cells in the channel catfish anterior kidney — consistent with its being a lymphohematopoietic tissue — but a relatively small number of AID-positive cells, which are closely associated with melanomacrophage cells.

Among the channel catfish tissues examined for AID mRNA expression, the highest amount of expression was observed in the spleen. This differs from the mouse and human whose spleens exhibited the lowest amount of AID mRNA expression among the secondary lymphoid tissues analyzed, with expression levels that were typically comparable to some primary lymphoid tissue

[Muramatsu, 1999] and some non-lymphoid tissues [Muto, 2000]. But, it was only various lymph nodes — structures that are absent in fish — that expressed more AID mRNA than the spleen in the mouse and human [Muramatsu, 1999; Muto, 2000; Saunders, 2004]. The expression pattern in adults of the frog *Xenopus*, which lack lymph nodes, was more similar to that of fish, with the highest amount of AID transcript found in the spleen and low levels of expression detectable in recognized MALT-containing tissues and select non-lymphoid tissues at higher numbers (35x) of PCR cycles [Ichikawa, 2006; Marr, 2007].

The tissue mRNA analysis revealed which tissues harbour AID-expressing cells in preparation for the more thorough histological examination necessary to resolve whether channel catfish possess an identifiable tissue architecture that might operate in lieu of the germinal centers found in higher vertebrates. Because clusters of proliferating and apoptotic cells are morphological signatures associated with germinal centers in mammals, patterns of cell proliferation [Oko, 2009] and death were examined to attempt to roughly locate sites of potential germinal center analogues in channel catfish spleen, an AID-expressing tissue. These investigations revealed that there are both proliferating and apoptotic cells diffusely scattered throughout the organ. But they did not reveal any telltale clusterings of proliferating and/or apoptotic cells, even in the vicinity of the antigen-trapping melanomacrophage centers, which may also operate as tingible body macrophage counterparts [Agius, 2003]. This suggests that these activities may not be strongly associated with the melanomacrophage centers, a finding that would be consistent with *Xenopus* where proliferating antigen-specific B-cells were found to be dispersed throughout the spleen rather than clustered [Hsu, 1998; Marr, 2007]. However, it is also possible that proliferating and/or apoptotic cells were concentrated in the vicinity of melanomacrophage centers but were not observed because the strong autofluorescence of the melanomacrophage cells masked detection of the fluor used in these assays (see Results).

It should be noted that the cells labeled for proliferation or death were not identified by type, so it cannot be concluded that any of them are specifically B-

cells. In fact, because ISH for IgH-expressing cells (discussed below) indicated that they mostly occur in white pulp, but the TUNEL cell death assay indicated that cell death occurs in both red and white pulp, many of the apoptotic cells are quite likely to be red blood cells (i.e. erythrocytes), which are nucleated in fish and therefore detectable by TUNEL cell death assay. The fish spleen, like that of higher vertebrates, serves as a site of recycling for senescent red blood cells through phagocytosis by macrophages [Fange, 1985; Tort Bardolet, 2003]. In fact, this activity is believed to be the primary source of the haemosiderin pigment found in the melanomacrophage cells of the spleen [Fange, 1985]. Moreover, that no TUNEL-labeling was observed in the melanomacrophage cells, which should contain phagocytosed red blood cells (i.e. degraded DNA), further supports the concern that fluor detection in the vicinity of the melanomacrophage centers was masked by autofluorescence of their pigments.

That a few secondary repertoire lineages have been identified from teleost spleens indicates that some amount of proliferation of B-cells does occur in the channel catfish spleen [Marianes, 2011; Yang, 2006]. But how much, if any, death of B-cells may be occurring there is unclear. As noted above (see Introduction), in mice, when apoptosis is experimentally blocked in B-cells undergoing antibody affinity maturation, the effect of positive selection is diluted and the overall affinity of the response is significantly reduced [Tarlinton, 2000]. Thus, the apparently weak selection in teleosts could be largely attributable to a lack of apoptosis if it truly does not play a role in antibody affinity maturation in teleosts or if other factors conspire to diminish the extent of its occurrence relative to that observed in the mouse and human.

While the channel catfish spleen, which is more diffusely organized than its mammalian counterpart [Fange, 1985], does not appear to exhibit signs of some general morphological signatures of the conventional germinal centers of mammals, the ISH revealed that it does have a consistent association between cells expressing AID mRNA and melanomacrophage cells [Saunders, 2010], the proposed teleost morphological analogues of the mammalian antigen-trapping FDCs [Fange, 1985]. This suggests that at least some sort of functional

aggregations occur. More specifically, ISH in spleen sections from immunized channel catfish revealed that AID-positive cells were located in the white pulp and typically occurred in clusters associated with melanomacrophage cells [Saunders, 2010].

Although there is limited space in the white pulp, it is unlikely that the apparent association between AID-positive cells and melanomacrophage cells was falsely reinforced by space constraints as there was ample area that was unoccupied by AID-positive cells. This was corroborated by RT-PCR for AID mRNA from samples of channel catfish spleen, obtained by laser-capture microdissection, which showed that expression of AID co-occurs principally with melanomacrophage centers with little or no expression in the residual white pulp [Oko, 2009; Saunders, 2010].

Sorting of the autofluorescent melanomacrophage cells from channel catfish spleen by FACS followed by RT-PCR for colony stimulating factor 1 receptor (CSF1-R) mRNA — a marker of macrophage identity — and AID mRNA confirmed that the melanomacrophage cells do not express AID [Oko, 2009; Saunders, 2010]. Thus, it is other distinct cells within the melanomacrophage centers that are expressing AID.

Although the aggregations of AID-positive cells and melanomacrophage cells did not appear to exhibit a single consistent organization, there were some more commonly observed arrangements such as AID-positive cells and melanomacrophage cells forming a sphere with the melanomacrophage cells toward the exterior and the AID-positive cells within the interior (Figure 14). While it is tempting to view this apparent cellular segregation as being somewhat akin to mammalian germinal center organization, with its inner core of AID-expressing cells surrounded by an outer ring containing FDC analogues loosely reminiscent of dark and light zone equivalents, respectively, other organizations were not uncommon in the channel catfish spleen as well as in the kidney and intestine. A more restricted view of the organization might become apparent if the sample sizes and runs of serial sections were increased, but it is possible that these associations have not evolved a level of definite organization. In fact, as noted

above (see Introduction), mammalian germinal centers do not strictly adhere to the standard egg-shaped organization and may vary based on several factors including tissue type [Allen, 2004; Allen, 2007; Batista, 2000; Manser, 2004; Paus, 2006; Tarlinton, 1998].

Although it was anticipated that the antigen-trapping melanomacrophage cells would likely be found to be associated with AID-expressing cells, the compact and defined organization of the association was somewhat unexpected (see Introduction). Even though a teleost analogue of the mammalian germinal center was being sought, it was expected that any such analogue would be more loosely organized. That is, given the apparently limited antibody affinity maturation in teleosts and absence of genuine FDCs, it was expected that a teleost analogue of the mammalian germinal center would be more likely to resemble the less-defined sites of antibody affinity maturation that occur in aberrant immune responses of mammals, such as the ectopic germinal centers observed in human autoimmunity (see Introduction)[Park, 2005]. Similarly, it was also recognized that sites of AID-expressing cells might bear no resemblance to mammalian germinal centers, a situation more similar to that of lymphotoxin-alpha-deficient mice, which do not form germinal centers and do not have detectable FDCs (see Introduction) [Matsumoto, 1997; Tarlinton, 1998]. Although the sites of antibody affinity maturation have not been identified in these mice, it does occur and involves SHM [Diaz, 2001].

Some melanomacrophage centers did not appear to be associated with AID-positive cells, even though they were close neighbours to melanomacrophage centers that did have AID-positive cells associated with them. If this observation is accurate (see below), two obvious explanations for the distinction present themselves immediately. In essence, the first explanation is spatial and the second explanation is functional. First, in mammalian germinal centers, there does not appear to be any trafficking of antibody affinity maturing B-cells between germinal centers. Despite clear differences in how the sites of antibody affinity maturing B-cells arise in mammals (e.g. FDC appear to develop in response to activated B-cells) and teleosts (e.g. melanomacrophage centers are

presumed to be pre-existing), this feature may be replicated in the teleosts. That is, once activated B-cells become associated with a particular melanomacrophage center, they then remain associated with it until they terminally differentiate. If so, to have evolved an adaptive immune strategy that curtails the response from occupying all possible sites of antibody affinity maturation suggests that it is more favourable to concentrate the main response participants. Concentration of the affinity maturing B-cells through retention at sites of initial melanomacrophage center association could have obvious benefits for antibody affinity improvement via selection (see Introduction). And an absence of trafficking of the B-cells between sites could improve the diversification of the response. But, the apparently limited antibody affinity increase observed in fish suggests that this strategy might be more about control of the response and possibly resource management.

The second explanation for the distinction between proximate but differentially AID-positive cell-associated melanomacrophage centers is that they may represent different functional subsets of melanomacrophage cells. Melanomacrophage cells have more functions (e.g. iron capture) than just those identified as potentially relevant to antibody affinity maturation (e.g. antigen-trapping). Some of these functions may be associated with distinct subsets of melanomacrophage cells or may merely render the cells temporarily unavailable to participate in other functions [Fange, 1985]. The multifunctional nature of melanomacrophage cells is well illustrated by contrasting the melanomacrophage cells of the spleen, which appear to play a significant role in antibody affinity maturation, with those of the liver, an organ in which very little expression of IgH mRNA and no expression of AID mRNA was detected [Fange, 1985; Saunders, 2010].

Additionally, it is possible that the accuracy of some of the observations from the ISH, such as the rare, apparently orphaned, AID-positive cells or proximate, but seemingly uninvolved, individual melanomacrophage cells, was influenced by the timing of the captured interactions (e.g. B-cells in transit to a melanomacrophage center) and/or are an artifact of the microscopy (e.g. not able

to assess z-dimension due to limited depth of sections and insufficient serial sections).

Intriguingly, analysis of adult *Xenopus* spleens by Marr et al. [Marr, 2007] indicated a looser distribution of AID-positive cells than that observed in both the channel catfish and mammals. Additionally, although the AID-positive cells were predominantly located in the splenic white pulp of adult *Xenopus*, they were also found in the splenic red pulp. Strangely, they indicate that the AID-positive cells observed in the red pulp appear morphologically to be secretory (i.e. “large cells with noticeable cytoplasm”) [Marr, 2007]. They note that in the immunized specimens, although the AID-positive cells sometimes appeared to be localized to the blood vessels, there was no apparent clustering or other discernable organization, a finding consistent with their observation that proliferating antigen-specific B-cells exhibited a “scattered distribution” [Marr, 2007]. Given that *Xenopus* seem to generate greater antibody affinity increase than teleosts, these findings suggest that a specialized germinal center-like site may not be as essential to the operation as it has been supposed to be from mammalian studies. It also reinforces that the colocalization of clusters of AID-positive cells and melanomacrophage cells observed in channel catfish likely represent germinal center analogues rather than primordial germinal centers.

In vitro, co-expression of both IgH mRNA and AID mRNA was found to occur in the PMA/CI stimulated 1B10 B-cells (Figure 20) [Saunders, 2010]. But, it initially appeared that, other than being restricted to the splenic white pulp, the *in vivo* locations of the AID-positive cells did not overlap with the locations of the IgH-positive cells. It was later determined that AID and IgH are both expressed in the melanomacrophage centers and the early confusion was the result of noted difficulties (e.g. sensitivity; see Results) with the NBT/BCIP detection method that was used for IgH ISH coupled with the heavy pigmentation and autofluorescence of the melanomacrophage cells.

Laser-capture microdissection of channel catfish spleen and semi-quantitative RT-PCR for IgH mRNA, a far more sensitive technique than ISH, indicated that IgH mRNA expression occurs within melanomacrophage centers to

approximately the same extent as it does within an area of equivalent size from the residual white pulp [Saunders, 2010]. Because ISH indicated that IgH-positive cells were scattered in the residual white pulp, the RT-PCR finding that IgH mRNA expression there is approximately equivalent to that within the melanomacrophage centers, where AID-positive cells congregate, suggests that level of IgH mRNA expression may be lower in the B-cells that associate with the melanomacrophage centers than in those that do not. If so, it may be that the channel catfish AID-positive cells have downregulated the expression of IgH, as has been suggested to occur in centroblasts in the mammalian paradigm of antibody affinity maturation (see Introduction).

With confirmation that both IgH and AID expression colocalized with the sites of the melanomacrophage centers, the ISH protocol and detection apparatus were altered to attempt to improve detection and resolution of the colocalized IgH signal within melanomacrophage centers. Hybridized IgH probe was detected via FITC-conjugated secondary antibody, instead of the combination of alkaline phosphatase-conjugated anti-DIG Fab fragments and NBT/BCIP, to eliminate the probe-independent deposition of coloured precipitate within the melanomacrophage centers. Confocal laser scanning microscopy, a far more sensitive method of visualization, was employed to unmix the epifluorescence of the FITC ISH signals from the autofluorescence of the melanomacrophage cells in channel catfish spleen sections [Saunders, 2010].

With this refined approach, it was confirmed that IgH-positive cells are in fact present within the melanomacrophage centers as distinct entities [Saunders, 2010]. That such a highly sensitive detection method was required to identify these melanomacrophage center-associated IgH-positive cells favours the conclusion that these cells express a lower level of IgH mRNA than the IgH-positive cells found scattered in the splenic white pulp as these latter cells stained so intensely with NBT/BCIP that the pigmentation of melanomacrophage cells and background staining is unlikely to have been sufficient to mask their detection when analyzed with the compound microscope (Figure 14).

Unmixing of signals via confocal laser scanning microscopy was not performed for detection of other signals. As noted above, this raises some serious concerns about the accuracy of the findings from the cell proliferation and cell death assays — in which proliferating or apoptotic cells, respectively, were observed but were not found to be associated with melanomacrophage centers in the channel catfish spleen — as autofluorescence may have masked signals within the melanomacrophage centers.

Failure to detect ISH signals for T-cell markers precluded identification of the location of T-cells relative to the melanomacrophage centers (i.e. sites of AID-positive cells). It is possible that this was due to very low expression of the probe targets since the signal was not detectable even outside of melanomacrophage centers. Laser-capture microdissection of channel catfish spleen and semi-quantitative RT-PCR indicated that expression of TCRbeta mRNA was lower than AID expression in the melanomacrophage centers [Saunders, 2010]. Additionally, samples from the melanomacrophage centers and the residual splenic white pulp were found to express similar levels of TCRbeta mRNA and of CD4 mRNA which suggests that T-cells are no more numerous at sites of melanomacrophage centers than they are elsewhere in the splenic white pulp [Saunders, 2010].

As noted in the Introduction, in the mammalian paradigm of antibody affinity maturation, germinal center T-cells may be a limiting factor in B-cell selection since direct interaction with T-cells appears to be necessary to abort apoptosis of B-cells. If this model of B-cell competition for direct T-cell interaction is replicated in the apparent teleost germinal center analogue, relatively low numbers of T-cells should theoretically translate to enhanced antibody affinity increase due to the stringent competition. That this is not the case hints that direct T-cell interactions may not be explicitly required for survival of B-cells in this instance, a scenario that would be consistent with apoptosis apparently not playing a central role in antibody affinity maturation in teleosts. Further research will be required to resolve these issues.

In addition to the discovery of an AID homologue in fish, the understanding of some other pertinent aspects of teleost antibody affinity maturation has been further elucidated. Isolation of melanomacrophage cells and clusters from the goldfish (*Carassius auratus*) confirmed that these pigment-containing macrophages, which have been identified as tissue-resident macrophages, are phenotypically distinct from other tissue macrophages [Diaz-Satizabal, 2015]. Melanomacrophage cells isolated from goldfish spleen and kidney exhibited a considerable range of green and red autofluorescence (approximately 100-fold difference between highest and lowest autofluorescence), cell size, internal complexity, and cytochemical staining characteristics within each organ and also between the two organs [Diaz-Satizabal, 2015]. It is not known if any aspect of this heterogeneity correlates with distinct differences in melanomacrophage functions or reliably marks specific subpopulations. The *ex vivo* survival of isolated melanomacrophage cells in culture was short (less than 4 days) and was not improved by culture techniques that promote the survival of tissue macrophages [Diaz-Satizabal, 2015]. Survival was slightly prolonged (to more than three weeks) when melanomacrophages were isolated as clusters (i.e. mixed with other cell types) and surviving melanomacrophage cells appeared to be engaged in adherent contact with other cell types (e.g. fibroblast-like cells) [Diaz-Satizabal, 2015]. This suggests that they may require physical engagement with other cell types for maintenance, although there are other explanations for the apparent interactions. Although it is not uncommon for cell survival to be dependent on cellular contact, this apparent reliance of melanomacrophages on other cell types found within melanomacrophage clusters for survival calls to mind germinal center FDCs. As noted in the Introduction, FDCs, which are now believed to arise from vascular mural cells, don’t occur in the absence of lymphocytes [Kranich, 2016; Matsumoto, 1997; Park, 2005].

4.3 Update: Expanded Roles For AID And An Alternative Interpretation Of AID Expression In Channel Catfish Kidney

Since the completion of this research, some notable advancements have been made in the understanding of AID evolution, function, post-translational regulation, and targeting, although little additional progress – beyond phylogenetic analyses – has been made in specifically understanding teleost AID. For clarity, some of the more recent information was noted in the Introduction and at relevant points of the Discussion, while the remainder is detailed here. Current phylogenetic analyses suggest that the evolutionary origins of the AID/APOBEC family of cytidine deaminases found in vertebrates, which now includes AID, APOBEC1, APOBEC2, APOBEC3A-H, and APOBEC4, can be traced, along with other eukaryote deaminases, to the deaminase genes of the bacterial polymorphic toxin system [Barreto, 2011; Iyer, 2011; Krishnan, 2018]. Iyer et al. [Iyer, 2011] propose that several eukaryotic deaminases can trace their origins to independent acquisition of elements from this highly divergent superfamily, the members of which may have diverse roles in intraspecific resource competition, kin recognition, and as virulence factors. However, this interpretation may be merely an artifact of the methodology due to a combination of sparse sequence data and the considerable diversification of these molecules.

It appears that the ancestral gene of the AID/APOBEC family was acquired before the divergence of the vertebrates, as two AID orthologues, cytidine deaminase (CDA) 1 and CDA2, have been identified in agnathans [Barreto, 2011; Hirano, 2015; Krishnan, 2018; Kuraoka, 2018; Morales Poole, 2017; Patel, 2018]. These fascinating organisms possess an adaptive immune system that is distinct from that of gnathostomes but which relies on many analogous and/or homologous elements [Barreto, 2011; Morales Poole, 2017; Patel, 2018]. Lamprey lack Ig and TCR but utilize variable lymphocyte receptors (VLR), which contain many leucine-rich repeats (LRR), to recognize antigen [Hirano, 2015; Morales Poole, 2017; Patel, 2018]. Analogous to the germline organization of Ig

and TCR genes, the germline VLR genes are incomplete and are located proximate to several LRR elements [Hirano, 2015; Morales Poole, 2017; Patel, 2018]. Complete VLR genes are generated from the proximate LRRs through a gene conversion-like mechanism during the development of the lamprey lymphocyte-equivalent cells [Hirano, 2015; Morales Poole, 2017]. It has been hypothesized that CDA1 and CDA2 facilitate this process – a role that may be loosely likened to that played by RAG1/2 in V(D)J recombination in outcome (i.e. creation of a functional gene) but not specific activity. Interestingly, differential expression of these two genes suggests that CDA1 may serve to generate the VLRA and VLRC genes in the two types of cells that appear to be the lamprey equivalents of αβ and γδ T-cells (i.e. VLRs remain membrane-bound) and CDA2 may serve to generate the VLRB gene in the single apparent B-cell equivalent (i.e. VLRs are both membrane-bound and secreted) [Hirano, 2015; Morales Poole, 2017; Patel, 2018].

And some members of the AID/APOBEC family that had been previously thought to be confined to higher vertebrates appear to have evolved earlier. Homologues of most AID/APOBEC family members have now been identified in fish, although not necessarily in teleosts [Conticello, 2008; Krishnan, 2018; Patel, 2018]. For example, while homologues of APOBEC1 and an APOBEC3C have been identified in lungfish, only homologues of APOBEC2 and APOBEC4, which are known to be more ancient, have been identified in teleosts in addition to AID [Conticello, 2008; Krishnan, 2018; Patel, 2018]. In fact, at least some teleosts have been found to possess two APOBEC2 genes – attributed to retention after the teleost-specific genome duplication event – which appear to be coevolving [Conticello, 2008]. Much of this phylogenetic analysis work has only recently become possible because of the rapid expansion of available sequence data for an increasing diversity of lower vertebrate and invertebrate species. As the genomes of more Actinopterygian (ray-finned fish) species become available, it is reasonable to speculate that more AID/APOBEC family homologues could be identified in this very large and highly diverse class of vertebrates.

Subsequent to the completion of this thesis research, the channel catfish genome sequence became available [Liu, 2016]. As it is possible that channel catfish could harbour additional AID variants not identified in the expression studies of the channel catfish immortal B-cell line and tissues, the Ictalurus punctatus genome (taxid: 7998), which has undergone automated computational analysis, was BLAST searched. A BLASTP search using the protein sequence predicted from the channel catfish AID homologue (AY436507) as the query sequence identified “predicted: single-stranded DNA cytosine deaminase isoform X1” (Accession number XP 017325629). This predicted isoform utilizes an alternative exon 1 such that the protein sequence of the isoform differs only in the amino terminus prior to AID residue 7 (i.e. MKG instead of MSKLDS). This putative isoform was not identified in 5’ RACE from the catfish spleen and manual analysis of this region of the genome suggests that it is not likely to be expressed. But this would need to be definitively determined by RT-PCR with an isoform-specific forward primer.

There has been limited functional characterization of AID/APOBEC family homologues identified in lower vertebrates. It is possible that some will be found to serve different functions from those identified in mammals. In fact, the full range of functions of AID and most other AID/APOBEC family members found in mammals are still being discovered. In addition to its roles in antibody diversification (i.e. SHM, CSR, and gene conversion) of both the primary and secondary repertoires, it appears that AID may also contribute to epigenetic reprogramming of primordial germ cells of vertebrates [Moris, 2014; Villota-Herdoiza, 2013]. It is believed to accomplish this via demethylation of 5-methylcytosine (5-mC) – a signature of gene-silencing – in the genome [Bochtler, 2017; Moris, 2014; Villota-Herdoiza, 2013]. But there is considerable controversy regarding exactly how it accomplishes this and whether it acts globally or targets specific sequences (e.g. promoters). There is no evidence that AID can act directly as a demethylase and AID-mediated demethylation appears to be achieved indirectly through deamination [Bochtler, 2017]. An early model proposed that AID deaminates 5-mC to yield T and the resulting T:G mismatch is

then repaired with an unmethylated C [Moris, 2014; Villota-Herdoiza, 2013]. A more recent model proposed that AID may not even act directly on the 5-mC but deaminates proximate C and subsequent “regional” repair results in the exchange of C for 5-mC [Bochtler, 2017]. Although both agree on a requirement for AID and repair, the considerable difference between these two proposed models highlights the uncertainty and ongoing debate in this issue. Regardless of the specific details of demethylation, it appears that, in this context, AID activity indirectly leads to demethylation of 5-mC rather than to the mutational outcomes that are observed in contexts associated with antibody diversification events.

It has also not yet been determined how the distinctly different outcomes of demethylation versus hypermutation are achieved. It is conceivable that context-specific post-translational regulation of AID and/or differential expression or regulation of co-factors *in vivo* could influence AID activity at 5-mC differently during epigenetic reprogramming versus AID activity during antibody diversification. Indeed, that AID has also been implicated in alterations to epigenetic methylation patterns in germinal center B-cells suggests that demethylation-associated deamination activity, as with SHM- and CSR-associated deamination activities, may be regulated by distinct mechanisms [Dominguez, 2015].

Since AID has been implicated in demethylation in zebrafish embryos, Abdouni et al. [Abdouni, 2013] tested the ability of purified recombinant AID from four teleost fish species to deaminate 5-mC *in vitro* using “bubble” substrates in which the target C or 5-mC was situated in an AID hotspot motif (i.e. WRC). Human AID and most fish AIDs, including that of channel catfish, deaminated 5-mC far less efficiently than C on these substrates at their respective optimal temperatures (i.e. 37°C for human AID and 25°C for fish AIDs) [Abdouni, 2013]. But zebrafish AID was able to deaminate 5-mC considerably more efficiently than the other AIDs, although still not as efficiently as it deaminated C [Abdouni, 2013]. As expected from previous experiments, sequence context (i.e. various WRC and non-WRC) influenced AID deamination efficiency on the various substrates [Abdouni, 2013]. Intriguingly, the 5-mC deamination efficiency of

zebrafish AID was higher when the 5-mC was part of CpG than when it wasn't [Abdouni, 2013].

Analysis of a variety of mutant and domain exchanged AIDs suggests that the higher 5-mC deamination efficiency of zebrafish AID is not a result of their relatively higher catalytic activity but relies on other features of the molecule (e.g. differences in structural flexibility) [Abdouni, 2013]. Abdouni et al. [Abdouni, 2013] speculate that the higher 5-mC deamination efficiency of zebrafish AID, including at CpG, may reflect its apparently "more prominent role" in demethylation in zebrafish embryogenesis relative to that of mouse and human, which may compensate with redundant mechanisms [Abdouni, 2013]. However, this interpretation suggests that AID-mediated demethylation may not play as significant a role in the other fish species analyzed as the 5-mC deamination efficiency of their AIDs was more similar to that of human AID [Abdouni, 2013]. Of course, other mechanisms (e.g. cofactors) likely influence AID activity at 5-mC *in vivo*, so differences in 5-mC deamination efficiency among AIDs may not translate directly to differences in demethylation *in vivo*. Indeed, recall that in one of the AID-mediated demethylation models cited above, demethylation does not rely on the deamination of 5-mC [Bochtler, 2017]. Clarification of the *in vivo* impact of differences in 5-mC deamination efficiency among AIDs awaits better resolution of the actual role of AID in demethylation during embryogenesis and analysis of the extent of the role of AID in demethylation during embryogenesis of other fish species.

In another developmental role, but one which is more specific to immune function, AID is now also believed to contribute to the enforcement of central tolerance during B-cell development [Kuraoka, 2018]. Central tolerance mechanisms suppress the development of autoreactive B-cells through receptor editing, deletion (i.e. apoptosis), and/or anergy. Central tolerance is not absolute and it appears that a notable percentage of potentially autoreactive B-cells do enter the periphery [Kuraoka, 2018]. But its absence (e.g. through blocked apoptosis) is associated with a significantly increased incidence of autoimmunity,

which is also a feature of HIGM2 (i.e. AID defect) [Durandy, 2006; Durandy, 2005; Kuraoka, 2018; Minegishi, 2000; Revy, 2000].

It appears that AID helps to suppress the development of autoreactive B-cells through a deletion mechanism that activates Caspase 3 (i.e. apoptosis) [Kuraoka, 2018]. The actual function of AID in this context is still being investigated but it has been postulated that it may act by generating DNA damage, possibly in concert with RAG activity (e.g. ineffective receptor editing), that triggers apoptosis [Kuraoka, 2018].

To achieve central tolerance, it appears that expression of AID is required at the immature/T1 B-cell stage (i.e. when the complete BCR is expressed on the cell surface) [Kuraoka, 2018]. Recognition of endogenous antigen (i.e. self) by autoreactive BCR leads to its endocytosis and transport to intracellular compartments containing endosomal TLRs (e.g. TLR 9) which can recognize damage-associated molecular patterns (i.e. DAMPs) [Kuraoka, 2018]. These DAMPs may be intrinsic to the endogenous antigen itself or to molecules that were endocytosed with it. Synergistic signaling through both the BCR and endosomal TLRs appears to be required to upregulate AID expression at this stage [Kuraoka, 2018]. While exogenous antigen, cognate T-cell help, and/or T-cells do not appear to be required for AID expression in this context, other extracellular factors may play a role in regulating the level of AID expression [Kuraoka, 2018]. For example, IL-7 – a cytokine produced by bone marrow stromal cells – may damp AID expression somewhat at this stage as IL-7R defects that abrogate response to IL-7 correlate with a higher level of AID expression [Kuraoka, 2018].

Of particular note here, it appears that signaling through TACI, which upregulates AID expression and leads to isotype switching in T-cell-independent responses to exogenous antigen (see Introduction), also plays a role in B-cell central tolerance, which is a T-cell-independent response to endogenous antigen [Garcia-Carmona, 2018]. Furthermore, TACI has been found to associate with TLR9 in the endosomal compartment [Garcia-Carmona, 2018]. Precisely how B-cells regulate differential response (i.e. isotype switch versus apoptosis) to the

recognition of exogenous or endogenous antigens, despite relying on many of the same elements (e.g. BCR, TACI system, TLRs, DAMPs, and AID) in both responses, remains to be determined.

The finding that AID expression is required to enforce central tolerance during B-cell development in mouse/human further complicates the interpretation of the AID expression observed in the anterior kidney of the channel catfish. As discussed above, although secondary antibody repertoire diversification was believed to be the most likely explanation for the AID expression observed in the anterior kidney, primary antibody repertoire diversification could not be absolutely ruled out. However, in light of the discovery of its role in central tolerance in mouse/human, it is reasonable to speculate that AID expression in the anterior kidney of the channel catfish could also be attributed to a role in central tolerance during the development of fish B-cells. Given the importance of central tolerance to suppression of autoimmunity in mammals and that there is no evidence that fish are prone to autoimmunity despite facing a similar risk for development of autoreactive B-cells, it is plausible that some form of central tolerance operates in fish. For the sake of discussion herein, the postulated central tolerance of fish will be presumed to rely on AID expression, but the occurrence of central tolerance mechanisms that don't rely on AID expression in these highly divergent organisms cannot be discounted.

The speculation that AID expression in channel catfish anterior kidney could be contributing to central tolerance assumes conservation of features beyond those required for SHM, as it is probable that at least some features (e.g. gene expression and protein-protein interactions) required for central tolerance differ from those required for SHM. This likely also applies to features required for its role in epigenetic reprogramming and, as discussed above, has been found to be the case for CSR function. In mammals, CSR function relies in part on features of AID that are distinct from those required for SHM [Shinkura, 2004]. Even though teleosts lack the genetic capacity for CSR, their AID homologue exhibits conservation of CSR function [Barreto, 2005; Wakae, 2006]. It is likely that this is partly due to the overlap of the CSR functional region with the NES sequence as

the need to conserve the NES consensus sequence places constraints on the evolution of the region. Despite the regional overlap of CSR and NES-mediated export functions, a study of carboxy-terminal mutant AIDs has indicated that there is a disconnect between export and CSR functions [Geisberger, 2009]. That is, mutations in the NES differentially impact export and CSR functions [Geisberger, 2009]. It will be interesting to learn which features of AID central tolerance relies on and how well conserved they are in teleosts.

And still other activities have been postulated for AID. For example, it has been suggested that AID may play a role in antiviral responses beyond its contributions to antibody affinity maturation but which is still indirect (i.e. unlike the APOBEC3 family molecules which directly deaminate viral nucleic acids) [Moris, 2014]. Although mechanisms that control AID gene expression and transcript regulation are beyond the immediate scope of this research and so have been largely omitted from this thesis, it is of note here that a variety of microRNAs (e.g. microRNAs 155, 181b, and 93) appear to contribute to post-transcriptional regulation of AID [Zan, 2013]. Intriguingly, it appears that some viruses may exploit the use of microRNA to reduce AID expression in host cells [Moris, 2014]. This seems to support the idea that AID may contribute to intracellular antiviral responses, as merely suppressing AID in infected cells is unlikely to have a significant impact on affinity maturation of antiviral antibodies unless viral infection involves rapid and pervasive infection of B-cells.

Given its multiple proposed roles in epigenetic reprogramming, antigen receptor diversification, and tolerance, it is no surprise that AID appears to be undergoing purifying selection even though some members of the AID/APOBEC family, such as the APOBEC3s, appear to be rapidly diversifying as a result of positive selection in response to evolving targets (e.g. viruses) [Krishnan, 2018]. Despite a few minor differences in sequence and activity, it appears that AID is relatively well conserved in vertebrates including structural features that regulate DNA-binding and access to the catalytic pocket [Quinlan, 2017]. Conservation among cis-acting genetic elements and trans-acting cellular factors and their

differential impact on AID-mediated functions in the vertebrate lineage remain to be determined.

5. Conclusion

This research determined that channel catfish express a functional AID homologue within B-cells in a tissue pattern that is consistent with an immune role. That mutation was found to have occurred in the endogenous IgH gene in a channel catfish immortal B-cell line subsequent to upregulation of native AID confirmed that channel catfish AID is catalytically active and that the channel catfish IgH gene organization supports SHM. And, in sections from channel catfish spleen and certain other tissues that express AID mRNA, *in situ* hybridization revealed that the cells that express AID mRNA colocalize in discrete clusters with the antigen-trapping melanomacrophage cells, which are putative teleost FDC analogues. That there appears to be a definite association between AID-positive cells and melanomacrophage cells suggests that these cellular aggregations likely have a functional purpose, which is postulated to be analogous to that of mammalian germinal centers. However, the precise nature of any functional interactions remains to be determined.

That the channel catfish is SHM-competent and appears to exhibit a tissue architecture that is remarkably reminiscent of mammalian germinal centers strongly implies the presence of selection. And recent genetic studies in teleosts have identified the presence of hypermutated clonal lineages. Yet the apparent inability of teleosts to significantly increase their antibody affinity, despite their genetic potential to achieve increases comparable to those observed in mammals, indicates that selection may be relatively inefficient. This conclusion is supported by the apparent absence in channel catfish of other attendant features of the intense selection in mammalian germinal centers, such as extensive cell death.

6. Significance

The findings from this thesis research were the basis for several additional projects designed to corroborate them and to better define the contributions of AID-mediated SHM and melanomacrophage cells to antibody affinity maturation in teleosts. These projects, some of which were noted in the Discussion, include:

1) The identification of the AID gene sequence of the channel catfish (cloning) [Saunders, 2004] enabled the study of the transcriptional regulation of the AID gene from channel catfish and zebrafish by transcription reporter assay in a variety of fish and mammalian cell types, including B-cells and non-B-cells [Villota-Herdoiza, 2013]. This work revealed that the transcriptional regulation of the teleost AID gene is complex. The activity of the teleost AID gene promoter, which was as robust as the SV40 control, was not restricted to B-cells. Testing of putative teleost AID gene regulatory regions in complex with the promoter revealed that the teleost AID gene contains regions that individually can exert a suppressor effect which imposes B-cell specificity on the promoter or that in combination can enhance transcription from the promoter in B-cells. This regulatory complexity and the potential for transcriptional activity in non-B-cells is consistent with the additional roles identified for AID which were described in the Discussion.
2) The identification of the AID gene sequence of the channel catfish (cloning) [Saunders, 2004] also contributed to analysis of the evolution of regulatory mechanisms that influence the subcellular localization of AID [Methot, 2015]. This work, which focused on elucidating the contributions of eEF1A to the cytoplasmic retention of AID, revealed that channel catfish AID is sequestered within channel catfish immortal B-cell line 1B10 cells in a pattern that is consistent with mammalian AID sequestration in mammalian B-cells: cytoplasmic retention and export from the nucleus. Thus, subcellular sequestration as a mechanism to regulate access of this potentially toxic protein to its substrate is conserved from fish to mammals.

3) The identification of melanomacrophage centers, which are known to trap antigen, as sites of cells that express AID mRNA in the channel catfish (ISH) [Saunders, 2010] provided substantial support for the hypothesis that these sites may be analogous to the germinal centers of mammals. Based on this finding, further characterization of melanomacrophage centers was pursued in goldfish because their melanomacrophage centers can be isolated intact for *in vitro* culture and analysis [Diaz-Satizabal, 2015]. This research determined that although melanomacrophage cells express the macrophage marker CSF1-R, they exhibit other phenotypic differences (e.g. autofluorescence, cytochemistry, and culture requirements) that suggest they are tissue-resident macrophages, a macrophage subset with origins distinct from macrophages that arise from immigrant monocytes. FACS analysis of melanomacrophage cells from the spleen and kidney, gated as sub-populations based on the relative intensity of their red and green autofluorescence, revealed that melanomacrophage cells exhibit considerable heterogeneity in autofluorescence, size, and complexity. Notably, heterogeneity of melanomacrophage cells was observed not only between the two organs, with the kidney harbouring a green-shifted sub-population that appears to be distinct from the splenic melanomacrophage cells, but also within each organ. Additional characterization of the role of melanomacrophage cells in antibody affinity maturation in teleosts is ongoing.
4) The demonstration that channel catfish AID is functional in a catfish B-cell line (mitogenic stimulation and mutation analysis) [Dancyger, 2012] coupled with the identification of melanomacrophage centers as sites of cells that express AID in channel catfish (ISH) [Saunders, 2010] marked melanomacrophage centers as target sites for SHM analyses. Zebrafish were chosen for subsequent studies of *in vivo* SHM and antibody affinity maturation because of the greater availability of genomic sequence. Doaa Waly (not yet published) obtained and analyzed the VDJ sequences of IgM and IgZ from the melanomacrophage centers of vaccinated and

unvaccinated zebrafish. She has determined that melanomacrophage centers are indeed sites of AID-mediated SHM in teleost fish. As in mammals, it appears that a very small number of B-cells nucleate the response in melanomacrophage centers and an additional small number of B-cells may immigrate later. Antigen-dependent repertoire shift and CDR3 clonal lineages were observed. Moreover, replacement mutations occurred more frequently within CDRs and less frequently in FR regions. This is consistent with SHM and selection data from mammals. This work, which is still in progress, will be followed up with antibody affinity analyses which are expected to corroborate that melanomacrophage centers are sites of antibody affinity maturation in fish.

References

1. Abdouni, H., J. J. King, M. Suliman, M. Quinlan, H. Fifield, and M. Larijani. 2013. Zebrafish AID is capable of deaminating methylated deoxycytidines. *Nucleic Acids Res* 41:5457-5468.
2. Agius, C., and R. J. Roberts. 2003. Melano-macrophage centres and their role in fish pathology. *J Fish Dis* 26:499-509.
3. Allen, C. D., K. M. Ansel, C. Low, R. Lesley, H. Tamamura, N. Fujii, and J. G. Cyster. 2004. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat Immunol* 5:943-952.
4. Allen, C. D., T. Okada, and J. G. Cyster. 2007. Germinal-center organization and cellular dynamics. *Immunity* 27:190-202.
5. Alrefai, R. H., D. B. Winter, V. A. Bohr, and P. J. Gearhart. 2007. Nucleotide excision repair in an immunoglobulin variable gene is less efficient than in a housekeeping gene. *Mol Immunol* 44:2800-2805.
6. Anderson, S. M., L. G. Hannum, and M. J. Shlomchik. 2006. Memory B cell survival and function in the absence of secreted antibody and immune complexes on follicular dendritic cells. *J Immunol* 176:4515-4519.
7. Aoufouchi, S., A. Faili, C. Zober, O. D'Orlando, S. Weller, J. C. Weill, and C. A. Reynaud. 2008. Proteasomal degradation restricts the nuclear lifespan of AID. *J Exp Med* 205:1357-1368.
8. Arakawa, H., J. Hauschild, and J. M. Buerstedde. 2002. Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. *Science* 295:1301-1306.
9. Bachl, J., C. Carlson, V. Gray-Schopfer, M. Dessing, and C. Olsson. 2001. Increased transcription levels induce higher mutation rates in a hypermutating cell line. *J Immunol* 166:5051-5057.
10. Barreto, V., B. Reina-San-Martin, A. R. Ramiro, K. M. McBride, and M. C. Nussenzweig. 2003. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol Cell* 12:501-508.
11. Barreto, V. M., and B. G. Magor. 2011. Activation-induced cytidine deaminase structure and functions: a species comparative view. *Dev Comp Immunol* 35:991-1007.
12. Barreto, V. M., Q. Pan-Hammarstrom, Y. Zhao, L. Hammarstrom, Z. Misulovin, and M. C. Nussenzweig. 2005. AID from bony fish catalyzes class switch recombination. *J Exp Med* 202:733-738.
13. Barreto, V. M., A. R. Ramiro, and M. C. Nussenzweig. 2005. Activation-induced deaminase: controversies and open questions. *Trends Immunol* 26:90-96.
14. Bassing, C. H., W. Swat, and F. W. Alt. 2002. The mechanism and regulation of chromosomal V(D)J recombination. *Cell* 109 Suppl:S45-55.
15. Batista, F. D., and M. S. Neuberger. 2000. B cells extract and present immobilized antigen: implications for affinity discrimination. *EMBO J* 19:513-520.

16. Benson, M. J., L. D. Erickson, M. W. Gleeson, and R. J. Noelle. 2007. Affinity of antigen encounter and other early B-cell signals determine B-cell fate. *Curr Opin Immunol* 19:275-280.
17. Besmer, E., P. Gourzi, and F. N. Papavasiliou. 2004. The regulation of somatic hypermutation. *Curr Opin Immunol* 16:241-245.
18. Binz, S. K., A. M. Sheehan, and M. S. Wold. 2004. Replication protein A phosphorylation and the cellular response to DNA damage. *DNA Repair (Amst)* 3:1015-1024.
19. Blumberg, P. M. 1988. Protein kinase C as the receptor for the phorbol ester tumor promoters: sixth Rhoads memorial award lecture. *Cancer Res* 48:1-8.
20. Bly, J. E., M. A. Cuchens, and L. W. Clem. 1986. Temperature-mediated processes in teleost immunity: binding and mitogenic properties of concanavalin A with channel catfish lymphocytes. *Immunology* 58:523-526.
21. Bochtler, M., A. Kolano, and G. L. Xu. 2017. DNA demethylation pathways: Additional players and regulators. *Bioessays* 39:1-13.
22. Braasch, I., M. Schartl, and J. N. Volff. 2008. The evolution of teleost pigmentation and the fish-specific genome duplication. *Journal of Fish Biology* 73:1891-1918.
23. Bransteitter, R., P. Pham, P. Calabrese, and M. F. Goodman. 2004. Biochemical analysis of hypermutational targeting by wild type and mutant activation-induced cytidine deaminase. *J Biol Chem* 279:51612-51621.
24. Bransteitter, R., P. Pham, M. D. Scharff, and M. F. Goodman. 2003. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci U S A* 100:4102-4107.
25. Bransteitter, R., J. L. Sneeden, S. Allen, P. Pham, and M. F. Goodman. 2006. First AID (activation-induced cytidine deaminase) is needed to produce high affinity isotype-switched antibodies. *J Biol Chem* 281:16833-16836.
26. Brar, S. S., M. Watson, and M. Diaz. 2004. Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. *J Biol Chem* 279:26395-26401.
27. Budzko, L., P. Jackowiak, K. Kamel, J. Sarzynska, J. M. Bujnicki, and M. Figlerowicz. 2017. Mutations in human AID differentially affect its ability to deaminate cytidine and 5-methylcytidine in ssDNA substrates in vitro. *Sci Rep* 7:3873.
28. Burton, R. F. 2002. Evolutionary determinants of normal arterial plasma pH in ectothermic vertebrates. *J Exp Biol* 205:641-650.
29. Cain, K. D., D. R. Jones, and R. L. Raison. 2002. Antibody-antigen kinetics following immunization of rainbow trout (Oncorhynchus mykiss) with a T-cell dependent antigen. *Dev Comp Immunol* 26:181-190.
30. Cascalho, M. 2004. Advantages and disadvantages of cytidine deamination. *J Immunol* 172:6513-6518.

31. Chaudhuri, J., C. Khuong, and F. W. Alt. 2004. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* 430:992-998.
32. Chaudhuri, J., M. Tian, C. Khuong, K. Chua, E. Pinaud, and F. W. Alt. 2003. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* 422:726-730.
33. Clark, L. A., S. Ganesan, S. Papp, and H. W. van Vlijmen. 2006. Trends in antibody sequence changes during the somatic hypermutation process. *J Immunol* 177:333-340.
34. Clem, L. W., J. E. Bly, C. F. Ellsaesser, C. J. Lobb, and N. W. Miller. 1990. Channel catfish as an unconventional model for immunological studies. *J Exp Zool Suppl* 4:123-125.
35. Clem, L. W., J. E. Bly, M. Wilson, V. G. Chinchar, T. Stuge, K. Barker, C. Luft, M. Rycyzyn, R. J. Hogan, T. van Lopik, and N. W. Miller. 1996. Fish immunology: the utility of immortalized lymphoid cells--a mini review. *Vet Immunol Immunopathol* 54:137-144.
36. Clem, L. W., E. Faulmann, N. W. Miller, C. Ellsaesser, C. J. Lobb, and M. A. Cuchens. 1984. Temperature-mediated processes in teleost immunity: differential effects of in vitro and in vivo temperatures on mitogenic responses of channel catfish lymphocytes. *Dev Comp Immunol* 8:313-322.
37. Conticello, S. G. 2008. The AID/APOBEC family of nucleic acid mutators. *Genome Biol* 9:229.
38. Conticello, S. G., C. J. Thomas, S. K. Petersen-Mahrt, and M. S. Neuberger. 2005. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. *Mol Biol Evol* 22:367-377.
39. Dancyger, A. M., J. J. King, M. J. Quinlan, H. Fifield, S. Tucker, H. L. Saunders, M. Berru, B. G. Magor, A. Martin, and M. Larijani. 2012. Differences in the enzymatic efficiency of human and bony fish AID are mediated by a single residue in the C terminus modulating single-stranded DNA binding. *FASEB J* 26:1517-1525.
40. Dedeoglu, F., B. Horwitz, J. Chaudhuri, F. W. Alt, and R. S. Geha. 2004. Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NFkappaB. *Int Immunol* 16:395-404.
41. Di Noia, J., and M. S. Neuberger. 2002. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* 419:43-48.
42. Diaz, M., and P. Casali. 2002. Somatic immunoglobulin hypermutation. *Curr Opin Immunol* 14:235-240.
43. Diaz, M., and M. F. Flajnik. 1998. Evolution of somatic hypermutation and gene conversion in adaptive immunity. *Immunol Rev* 162:13-24.
44. Diaz, M., M. F. Flajnik, and N. Klinman. 2001. Evolution and the molecular basis of somatic hypermutation of antigen receptor genes. *Philos Trans R Soc Lond B Biol Sci* 356:67-72.

45. Diaz-Satizabal, L., and B. G. Magor. 2015. Isolation and cytochemical characterization of melanomacrophages and melanomacrophage clusters from goldfish (Carassius auratus, L.). *Dev Comp Immunol* 48:221-228.
46. Dickerson, S. K., E. Market, E. Besmer, and F. N. Papavasiliou. 2003. AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med* 197:1291-1296.
47. Djurdjevic, I., M. E. Kreft, and S. Susnik Bajec. 2015. Comparison of pigment cell ultrastructure and organisation in the dermis of marble trout and brown trout, and first description of erythrophore ultrastructure in salmonids. *J Anat* 227:583-595.
48. Dominguez, P. M., M. Teater, N. Chambwe, M. Kormaksson, D. Redmond, J. Ishii, B. Vuong, J. Chaudhuri, A. Melnick, A. Vasanthakumar, L. A. Godley, F. N. Papavasiliou, O. Elemento, and R. Shaknovich. 2015. DNA Methylation Dynamics of Germinal Center B Cells Are Mediated by AID. *Cell Rep* 12:2086-2098.
49. Durandy, A., S. Peron, N. Taubenheim, and A. Fischer. 2006. Activation-induced cytidine deaminase: structure-function relationship as based on the study of mutants. *Hum Mutat* 27:1185-1191.
50. Durandy, A., P. Revy, K. Imai, and A. Fischer. 2005. Hyper-immunoglobulin M syndromes caused by intrinsic B-lymphocyte defects. *Immunol Rev* 203:67-79.
51. Edholm, E. S., E. Bengten, J. L. Stafford, M. Sahoo, E. B. Taylor, N. W. Miller, and M. Wilson. 2010. Identification of two IgD+ B cell populations in channel catfish, Ictalurus punctatus. *J Immunol* 185:4082-4094.
52. Fange, R., and S. Nilsson. 1985. The fish spleen: structure and function. *Experientia* 41:152-158.
53. Foote, J., and H. N. Eisen. 1995. Kinetic and affinity limits on antibodies produced during immune responses. *Proc Natl Acad Sci U S A* 92:1254-1256.
54. Forbes, A. A., T. H. Powell, N. F. Lobo, and J. L. Feder. 2007. Resolving a DNA sequencing artifact associated with topoisomerase I generated clones in the plasmid pCR2.1. *Biotechniques* 42:458, 460, 462.
55. Freshney, R. I. 2000. *Culture of animal cells : a manual of basic technique*. J. Wiley, New York.
56. Fukita, Y., H. Jacobs, and K. Rajewsky. 1998. Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity* 9:105-114.
57. Ganesh, K., S. Adam, B. Taylor, P. Simpson, C. Rada, and M. Neuberger. 2011. CTNNBL1 is a novel nuclear localization sequence-binding protein that recognizes RNA-splicing factors CDC5L and Prp31. *J Biol Chem* 286:17091-17102.
58. Garcia-Carmona, Y., A. T. Ting, L. Radigan, S. K. Athuluri Divakar, J. Chavez, E. Meffre, A. Cerutti, and C. Cunningham-Rundles. 2018. TACI Isoforms Regulate Ligand Binding and Receptor Function. *Front Immunol* 9:2125.

59. Geisberger, R., C. Rada, and M. S. Neuberger. 2009. The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proc Natl Acad Sci U S A* 106:6736-6741.
60. Ginhoux, F., and M. Guilliams. 2016. Tissue-Resident Macrophage Ontogeny and Homeostasis. *Immunity* 44:439-449.
61. Gordon, M. S., C. M. Kanegai, J. R. Doerr, and R. Wall. 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.
62. Gramlich, H. S., T. Reisbig, and D. G. Schatz. 2012. AID-targeting and hypermutation of non-immunoglobulin genes does not correlate with proximity to immunoglobulin genes in germinal center B cells. *PLoS One* 7:e39601.
63. Green, N. S., M. M. Lin, and M. D. Scharff. 1998. Somatic hypermutation of antibody genes: a hot spot warms up. *Bioessays* 20:227-234.
64. Grizzle, J. M., and W. A. Rogers. 1976. *Anatomy and Histology of the Channel Catfish*. Alabama Agricultural Experiment Station. Auburn University, Alabama.
65. Grundstrom, C., A. Kumar, A. Priya, N. Negi, and T. Grundstrom. 2018. ETS1 and PAX5 transcription factors recruit AID to Igh DNA. *Eur J Immunol* 48:1687-1697.
66. Hannum, L. G., A. M. Haberman, S. M. Anderson, and M. J. Shlomchik. 2000. Germinal center initiation, variable gene region hypermutation, and mutant B cell selection without detectable immune complexes on follicular dendritic cells. *J Exp Med* 192:931-942.
67. Harris, R. S., S. K. Petersen-Mahrt, and M. S. Neuberger. 2002. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* 10:1247-1253.
68. Harris, R. S., J. E. Sale, S. K. Petersen-Mahrt, and M. S. Neuberger. 2002. AID is essential for immunoglobulin V gene conversion in a cultured B cell line. *Curr Biol* 12:435-438.
69. Higdon, C. W., R. D. Mitra, and S. L. Johnson. 2013. Gene expression analysis of zebrafish melanocytes, iridophores, and retinal pigmented epithelium reveals indicators of biological function and developmental origin. *PLoS One* 8:e67801.
70. Hirano, M. 2015. Evolution of vertebrate adaptive immunity: immune cells and tissues, and AID/APOBEC cytidine deaminases. *Bioessays* 37:877-887.
71. Honjo, T., M. Muramatsu, and S. Fagarasan. 2004. AID: how does it aid antibody diversity? *Immunity* 20:659-668.
72. Hsu, E. 1998. Mutation, selection, and memory in B lymphocytes of exothermic vertebrates. *Immunol Rev* 162:25-36.
73. Hsu, E. 2016. Assembly and Expression of Shark Ig Genes. *J Immunol* 196:3517-3523.
74. Hsu, E., N. Pulham, L. L. Rumfelt, and M. F. Flajnik. 2006. The plasticity of immunoglobulin gene systems in evolution. *Immunol Rev* 210:8-26.

75. Hurley, I. A., R. L. Mueller, K. A. Dunn, E. J. Schmidt, M. Friedman, R. K. Ho, V. E. Prince, Z. Yang, M. G. Thomas, and M. I. Coates. 2007. A new time-scale for ray-finned fish evolution. *Proc Biol Sci* 274:489-498.
76. Ichikawa, H. T., M. P. Sowden, A. T. Torelli, J. Bachl, P. Huang, G. S. Dance, S. H. Marr, J. Robert, J. E. Wedekind, H. C. Smith, and A. Bottaro. 2006. Structural phylogenetic analysis of activation-induced deaminase function. *J Immunol* 177:355-361.
77. Imai, K., G. Slupphaug, W. I. Lee, P. Revy, S. Nonoyama, N. Catalan, L. Yel, M. Forveille, B. Kavli, H. E. Krokan, H. D. Ochs, A. Fischer, and A. Durandy. 2003. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. *Nat Immunol* 4:1023-1028.
78. Ito, S., H. Nagaoka, R. Shinkura, N. Begum, M. Muramatsu, M. Nakata, and T. Honjo. 2004. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc Natl Acad Sci U S A* 101:1975-1980.
79. Iwama, G. K., and T. Nakanishi. 1996. *Fish physiology. Volume 15, Fish immune system: organism, pathogen, and environment*. Academic Press, New York.
80. Iyer, L. M., D. Zhang, I. B. Rogozin, and L. Aravind. 2011. Evolution of the deaminase fold and multiple origins of eukaryotic editing and mutagenic nucleic acid deaminases from bacterial toxin systems. *Nucleic Acids Res* 39:9473-9497.
81. Jacobs, H., and L. Bross. 2001. Towards an understanding of somatic hypermutation. *Curr Opin Immunol* 13:208-218.
82. Kaattari, S. L., H. L. Zhang, I. W. Khor, I. M. Kaattari, and D. A. Shapiro. 2002. Affinity maturation in trout: clonal dominance of high affinity antibodies late in the immune response. *Dev Comp Immunol* 26:191-200.
83. Kavli, B., S. Andersen, M. Otterlei, N. B. Liabakk, K. Imai, A. Fischer, A. Durandy, H. E. Krokan, and G. Slupphaug. 2005. B cells from hyper-IgM patients carrying UNG mutations lack ability to remove uracil from ssDNA and have elevated genomic uracil. *J Exp Med* 201:2011-2021.
84. Kavli, B., M. Otterlei, G. Slupphaug, and H. E. Krokan. 2007. Uracil in DNA--general mutagen, but normal intermediate in acquired immunity. *DNA Repair (Amst)* 6:505-516.
85. Kelsoe, G. 1996. Life and death in germinal centers (redux). *Immunity* 4:107-111.
86. Kenter, A. L., and P. Bhattacharya. 2004. AID: a very old motif newly recognized. *Nat Immunol* 5:1203-1204.
87. Kiernan, J. A. 1999. *Histological and histochemical methods : theory and practice*. Butterworth Heinemann, Oxford ; Boston.
88. King, J. J., and M. Larijani. 2017. A Novel Regulator of Activation-Induced Cytidine Deaminase/APOBECs in Immunity and Cancer: Schrodinger's CATalytic Pocket. *Front Immunol* 8:351.

89. Kinoshita, K., and T. Nonaka. 2006. The dark side of activation-induced cytidine deaminase: relationship with leukemia and beyond. *Int J Hematol* 83:201-207.
90. Kovtun, I. V., and C. T. McMurray. 2007. Crosstalk of DNA glycosylases with pathways other than base excision repair. *DNA Repair (Amst)* 6:517-529.
91. Kranich, J., and N. J. Krautler. 2016. How Follicular Dendritic Cells Shape the B-Cell Antigenome. *Front Immunol* 7:225.
92. Krappmann, D., A. Patke, V. Heissmeyer, and C. Scheidereit. 2001. B-cell receptor- and phorbol ester-induced NF-kappaB and c-Jun N-terminal kinase activation in B cells requires novel protein kinase C's. *Mol Cell Biol* 21:6640-6650.
93. Krishnan, A., L. M. Iyer, S. J. Holland, T. Boehm, and L. Aravind. 2018. Diversification of AID/APOBEC-like deaminases in metazoa: multiplicity of clades and widespread roles in immunity. *Proc Natl Acad Sci U S A* 115:E3201-E3210.
94. Kumar, S., and S. B. Hedges. 1998. A molecular timescale for vertebrate evolution. *Nature* 392:917-920.
95. Kuraoka, M., E. Meffre, and G. Kelsoe. 2018. The First B-Cell Tolerance Checkpoint in Mice and Humans: Control by AID. *Adv Immunol* 139:51-92.
96. Kuzin, II, J. E. Snyder, G. D. Ugine, D. Wu, S. Lee, T. Bushnell, Jr., R. A. Insel, F. M. Young, and A. Bottaro. 2001. Tetracyclines inhibit activated B cell function. *Int Immunol* 13:921-931.
97. Larijani, M., D. Frieder, T. M. Sonbuchner, R. Bransteitter, M. F. Goodman, E. E. Bouhassira, M. D. Scharff, and A. Martin. 2005. Methylation protects cytidines from AID-mediated deamination. *Mol Immunol* 42:599-604.
98. Lee, C. G., K. Kinoshita, A. Arudchandran, S. M. Cerritelli, R. J. Crouch, and T. Honjo. 2001. Quantitative regulation of class switch recombination by switch region transcription. *J Exp Med* 194:365-374.
99. Li, J., D. R. Barreda, Y. A. Zhang, H. Boshra, A. E. Gelman, S. Lapatra, L. Tort, and J. O. Sunyer. 2006. B lymphocytes from early vertebrates have potent phagocytic and microbicidal abilities. *Nat Immunol* 7:1116-1124.
100. Li, Z., C. J. Woo, and M. D. Scharff. 2003. Mutations in AID and UNG extend the function of AID. *Nat Immunol* 4:945-946.
101. Lin, G. L., C. F. Ellsaesser, L. W. Clem, and N. W. Miller. 1992. Phorbol ester/calcium ionophore activate fish leukocytes and induce long-term cultures. *Dev Comp Immunol* 16:153-163.
102. Lin, J. Y., and D. E. Fisher. 2007. Melanocyte biology and skin pigmentation. *Nature* 445:843-850.
103. Liu, Z., S. Liu, J. Yao, L. Bao, J. Zhang, Y. Li, C. Jiang, L. Sun, R. Wang, Y. Zhang, T. Zhou, Q. Zeng, Q. Fu, S. Gao, N. Li, S. Koren, Y. Jiang, A. Zimin, P. Xu, A. M. Phillippy, X. Geng, L. Song, F. Sun, C. Li, X. Wang, A. Chen, Y. Jin, Z. Yuan, Y. Yang, S. Tan, E. Peatman, J. Lu, Z. Qin, R. Dunham, Z. Li, T. Sonstegard, J. Feng, R. G. Danzmann, S. Schroeder, B. Scheffler, M. V. Duke, L. Ballard, H. Kucuktas, L. Kaltenboeck, H. Liu, J.

Armbruster, Y. Xie, M. L. Kirby, Y. Tian, M. E. Flanagan, W. Mu, and G. C. Waldbieser. 2016. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. *Nat Commun* 7:11757.
104. Lorin, T., F. G. Brunet, V. Laudet, and J. N. Volff. 2018. Teleost Fish-Specific Preferential Retention of Pigmentation Gene-Containing Families After Whole Genome Duplications in Vertebrates. *G3 (Bethesda)* 8:1795-1806.
105. Magor, B. G. 2015. Antibody Affinity Maturation in Fishes-Our Current Understanding. *Biology (Basel)* 4:512-524.
106. Magor, B. G., D. A. Ross, L. Pilstrom, and G. W. Warr. 1999. Transcriptional enhancers and the evolution of the IgH locus. *Immunol Today* 20:13-17.
107. Magor, B. G., M. R. Wilson, N. W. Miller, L. W. Clem, D. L. Middleton, and G. W. Warr. 1994. An Ig heavy chain enhancer of the channel catfish Ictalurus punctatus: evolutionary conservation of function but not structure. *J Immunol* 153:5556-5563.
108. Maizels, N. 2000. Secret sharers in the immune system: a novel RNA editing activity links switch recombination and somatic hypermutation. *Genome Biol* 1:REVIEWS1025.
109. Manser, T. 2004. Textbook germinal centers? *J Immunol* 172:3369-3375.
110. Marianes, A. E., and A. M. Zimmerman. 2011. Targets of somatic hypermutation within immunoglobulin light chain genes in zebrafish. *Immunology* 132:240-255.
111. Marr, S., H. Morales, A. Bottaro, M. Cooper, M. Flajnik, and J. Robert. 2007. Localization and differential expression of activation-induced cytidine deaminase in the amphibian Xenopus upon antigen stimulation and during early development. *J Immunol* 179:6783-6789.
112. Martin, A., P. D. Bardwell, C. J. Woo, M. Fan, M. J. Shulman, and M. D. Scharff. 2002. Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. *Nature* 415:802-806.
113. Matsumoto, M., Y. X. Fu, H. Molina, and D. D. Chaplin. 1997. Lymphotoxin-alpha-deficient and TNF receptor-I-deficient mice define developmental and functional characteristics of germinal centers. *Immunol Rev* 156:137-144.
114. McBride, K. M., V. Barreto, A. R. Ramiro, P. Stavropoulos, and M. C. Nussenzweig. 2004. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J Exp Med* 199:1235-1244.
115. McBride, K. M., A. Gazumyan, E. M. Woo, V. M. Barreto, D. F. Robbiani, B. T. Chait, and M. C. Nussenzweig. 2006. Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. *Proc Natl Acad Sci U S A* 103:8798-8803.
116. Meffre, E., N. Catalan, F. Seltz, A. Fischer, M. C. Nussenzweig, and A. Durandy. 2001. Somatic hypermutation shapes the antibody repertoire of memory B cells in humans. *J Exp Med* 194:375-378.

117. Methot, S. P., L. C. Litzler, F. Trajtenberg, A. Zahn, F. Robert, J. Pelletier, A. Buschiazzo, B. G. Magor, and J. M. Di Noia. 2015. Consecutive interactions with HSP90 and eEF1A underlie a functional maturation and storage pathway of AID in the cytoplasm. *J Exp Med* 212:581-596.
118. Michael, N., T. E. Martin, D. Nicolae, N. Kim, K. Padjen, P. Zhan, H. Nguyen, C. Pinkert, and U. Storb. 2002. Effects of sequence and structure on the hypermutability of immunoglobulin genes. *Immunity* 16:123-134.
119. Michael, N., H. M. Shen, S. Longerich, N. Kim, A. Longacre, and U. Storb. 2003. The E box motif CAGGTG enhances somatic hypermutation without enhancing transcription. *Immunity* 19:235-242.
120. Miller, N. W., V. G. Chinchar, and L. W. Clem. 1994. Development of leukocyte cell lines from the channel catfish (Ictalurus punctatus). *J Tissue Cult Meth* 16:117-123.
121. Miller, N. W., and L. W. Clem. 1984. Temperature-mediated processes in teleost immunity: differential effects of temperature on catfish in vitro antibody responses to thymus-dependent and thymus-independent antigens. *J Immunol* 133:2356-2359.
122. Miller, N. W., M. A. Rycyzyn, M. R. Wilson, G. W. Warr, J. P. Naftel, and L. W. Clem. 1994. Development and characterization of channel catfish long term B cell lines. *J Immunol* 152:2180-2189.
123. Milstein, C., M. S. Neuberger, and R. Staden. 1998. Both DNA strands of antibody genes are hypermutation targets. *Proc Natl Acad Sci U S A* 95:8791-8794.
124. Minegishi, Y., A. Lavoie, C. Cunningham-Rundles, P. M. Bedard, J. Hebert, L. Cote, K. Dan, D. Sedlak, R. H. Buckley, A. Fischer, A. Durandy, and M. E. Conley. 2000. Mutations in activation-induced cytidine deaminase in patients with hyper IgM syndrome. *Clin Immunol* 97:203-210.
125. Mitta, G., F. Vandenbulcke, T. Noel, B. Romestand, J. C. Beauvillain, M. Salzet, and P. Roch. 2000. Differential distribution and defence involvement of antimicrobial peptides in mussel. *J Cell Sci* 113 (Pt 15):2759-2769.
126. Morales Poole, J. R., J. Paganini, and P. Pontarotti. 2017. Convergent evolution of the adaptive immune response in jawed vertebrates and cyclostomes: An evolutionary biology approach based study. *Dev Comp Immunol* 75:120-126.
127. Moris, A., S. Murray, and S. Cardinaud. 2014. AID and APOBECs span the gap between innate and adaptive immunity. *Front Microbiol* 5:534.
128. Mort, R. L., I. J. Jackson, and E. E. Patton. 2015. The melanocyte lineage in development and disease. *Development* 142:620-632.
129. Moscat, J., M. T. Diaz-Meco, and P. Rennert. 2003. NF-kappaB activation by protein kinase C isoforms and B-cell function. *EMBO Rep* 4:31-36.
130. Mu, Y., and K. M. McBride. 2018. Targeting mutagenesis in B cells: Phosphorylation goes beyond AID association. *Mol Cell Oncol* 5:e1432259.

131. Muramatsu, M., K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai, and T. Honjo. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102:553-563.
132. Muramatsu, M., V. S. Sankaranand, S. Anant, M. Sugai, K. Kinoshita, N. O. Davidson, and T. Honjo. 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.
133. Muto, T., M. Muramatsu, M. Taniwaki, K. Kinoshita, and T. Honjo. 2000. Isolation, tissue distribution, and chromosomal localization of the human activation-induced cytidine deaminase (AID) gene. *Genomics* 68:85-88.
134. Nagaoka, H., M. Muramatsu, N. Yamamura, K. Kinoshita, and T. Honjo. 2002. Activation-induced deaminase (AID)-directed hypermutation in the immunoglobulin Smu region: implication of AID involvement in a common step of class switch recombination and somatic hypermutation. *J Exp Med* 195:529-534.
135. Nambu, Y., M. Sugai, H. Gonda, C. G. Lee, T. Katakai, Y. Agata, Y. Yokota, and A. Shimizu. 2003. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science* 302:2137-2140.
136. Neuberger, M. S., J. M. Di Noia, R. C. Beale, G. T. Williams, Z. Yang, and C. Rada. 2005. Somatic hypermutation at A.T pairs: polymerase error versus dUTP incorporation. *Nat Rev Immunol* 5:171-178.
137. Neuberger, M. S., R. S. Harris, J. Di Noia, and S. K. Petersen-Mahrt. 2003. Immunity through DNA deamination. *Trends Biochem Sci* 28:305-312.
138. Neuberger, M. S., and C. Milstein. 1995. Somatic hypermutation. *Curr Opin Immunol* 7:248-254.
139. Ohta, Y., and M. Flajnik. 2006. IgD, like IgM, is a primordial immunoglobulin class perpetuated in most jawed vertebrates. *Proc Natl Acad Sci U S A* 103:10723-10728.
140. Okazaki, I. M., H. Hiai, N. Kakazu, S. Yamada, M. Muramatsu, K. Kinoshita, and T. Honjo. 2003. Constitutive expression of AID leads to tumorigenesis. *J Exp Med* 197:1173-1181.
141. Okazaki, I. M., K. Kinoshita, M. Muramatsu, K. Yoshikawa, and T. Honjo. 2002. The AID enzyme induces class switch recombination in fibroblasts. *Nature* 416:340-345.
142. Oko, A. 2009. Characterization of melanomacrophage centres as germinal centre-like structures in the channel catfish, *Ictalurus punctatus*. In *Biological Sciences*. University of Alberta, Edmonton. 105.
143. Osawa, M. 2008. Melanocyte stem cells.
144. Osborne, C. S., L. Chakalova, K. E. Brown, D. Carter, A. Horton, E. Debrand, B. Goyenechea, J. A. Mitchell, S. Lopes, W. Reik, and P. Fraser. 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* 36:1065-1071.

145. Ott, J. A., C. D. Castro, T. C. Deiss, Y. Ohta, M. F. Flajnik, and M. F. Criscitiello. 2018. Somatic hypermutation of T cell receptor alpha chain contributes to selection in nurse shark thymus. *Elife* 7.
146. Papavasiliou, F. N., and D. G. Schatz. 2002. Somatic hypermutation of immunoglobulin genes: merging mechanisms for genetic diversity. *Cell* 109 Suppl:S35-44.
147. Park, C. S., and Y. S. Choi. 2005. How do follicular dendritic cells interact intimately with B cells in the germinal centre? *Immunology* 114:2-10.
148. Patel, B., R. Banerjee, M. Samanta, and S. Das. 2018. Diversity of Immunoglobulin (Ig) Isotypes and the Role of Activation-Induced Cytidine Deaminase (AID) in Fish. *Mol Biotechnol* 60:435-453.
149. Patenaude, A. M., A. Orthwein, Y. Hu, V. A. Campo, B. Kavli, A. Buschiazzo, and J. M. Di Noia. 2009. Active nuclear import and cytoplasmic retention of activation-induced deaminase. *Nat Struct Mol Biol* 16:517-527.
150. Paul, W. E. 1999. *Fundamental immunology*. Lippincott-Raven, Philadelphia.
151. Paus, D., T. G. Phan, T. D. Chan, S. Gardam, A. Basten, and R. Brink. 2006. Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. *J Exp Med* 203:1081-1091.
152. Peters, A., and U. Storb. 1996. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* 4:57-65.
153. Petersen-Mahrt, S. K., R. S. Harris, and M. S. Neuberger. 2002. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* 418:99-103.
154. Pham, P., R. Bransteitter, and M. F. Goodman. 2005. Reward versus risk: DNA cytidine deaminases triggering immunity and disease. *Biochemistry* 44:2703-2715.
155. Pham, P., R. Bransteitter, J. Petruska, and M. F. Goodman. 2003. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* 424:103-107.
156. Pilzecker, B., and H. Jacobs. 2019. Mutating for Good: DNA Damage Responses During Somatic Hypermutation. *Front Immunol* 10:438.
157. Poltoratsky, V., M. F. Goodman, and M. D. Scharff. 2000. Error-prone candidates vie for somatic mutation. *J Exp Med* 192:F27-30.
158. Poltoratsky, V., R. Prasad, J. K. Horton, and S. H. Wilson. 2007. Down-regulation of DNA polymerase beta accompanies somatic hypermutation in human BL2 cell lines. *DNA Repair (Amst)* 6:244-253.
159. Poltoratsky, V., C. J. Woo, B. Tippin, A. Martin, M. F. Goodman, and M. D. Scharff. 2001. Expression of error-prone polymerases in BL2 cells activated for Ig somatic hypermutation. *Proc Natl Acad Sci U S A* 98:7976-7981.
160. Poltoratsky, V. P., S. H. Wilson, T. A. Kunkel, and Y. I. Pavlov. 2004. Recombinogenic phenotype of human activation-induced cytosine deaminase. *J Immunol* 172:4308-4313.

161. Qin, X. F., A. Reichlin, Y. Luo, R. G. Roeder, and M. C. Nussenzweig. 1998. OCA-B integrates B cell antigen receptor-, CD40L- and IL 4-mediated signals for the germinal center pathway of B cell development. *EMBO J* 17:5066-5075.
162. Quinlan, E. M., J. J. King, C. T. Amemiya, E. Hsu, and M. Larijani. 2017. Biochemical Regulatory Features of Activation-Induced Cytidine Deaminase Remain Conserved from Lampreys to Humans. *Mol Cell Biol* 37.
163. Rada, C., J. M. Di Noia, and M. S. Neuberger. 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.
164. Rada, C., J. M. Jarvis, and C. Milstein. 2002. AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc Natl Acad Sci U S A* 99:7003-7008.
165. Rada, C., and C. Milstein. 2001. The intrinsic hypermutability of antibody heavy and light chain genes decays exponentially. *EMBO J* 20:4570-4576.
166. Rada, C., G. T. Williams, H. Nilsen, D. E. Barnes, T. Lindahl, and M. S. Neuberger. 2002. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr Biol* 12:1748-1755.
167. Ramiro, A. R., and M. C. Nussenzweig. 2004. Immunology: aid for AID. *Nature* 430:980-981.
168. Ramiro, A. R., P. Stavropoulos, M. Jankovic, and M. C. Nussenzweig. 2003. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat Immunol* 4:452-456.
169. Revy, P., T. Muto, Y. Levy, F. Geissmann, A. Plebani, O. Sanal, N. Catalan, M. Forveille, R. Dufourcq-Labelouse, A. Gennery, I. Tezcan, F. Ersoy, H. Kayserili, A. G. Ugazio, N. Brousse, M. Muramatsu, L. D. Notarangelo, K. Kinoshita, T. Honjo, A. Fischer, and A. Durandy. 2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* 102:565-575.
170. Rogozin, I. B., and M. Diaz. 2004. Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. *J Immunol* 172:3382-3384.
171. Santini, F., L. J. Harmon, G. Carnevale, and M. E. Alfaro. 2009. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes. *BMC Evol Biol* 9:194.
172. Saunders, H. L., and B. G. Magor. 2004. Cloning and expression of the AID gene in the channel catfish. *Dev Comp Immunol* 28:657-663.
173. Saunders, H. L., A. L. Oko, A. N. Scott, C. W. Fan, and B. G. Magor. 2010. The cellular context of AID expressing cells in fish lymphoid tissues. *Dev Comp Immunol* 34:669-676.

174. Shinkura, R., S. Ito, N. A. Begum, H. Nagaoka, M. Muramatsu, K. Kinoshita, Y. Sakakibara, H. Hijikata, and T. Honjo. 2004. Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat Immunol* 5:707-712.
175. Sizemore, R. C., N. W. Miller, M. A. Cuchens, C. J. Lobb, and L. W. Clem. 1984. Phylogeny of lymphocyte heterogeneity: the cellular requirements for in vitro mitogenic responses of channel catfish leukocytes. *J Immunol* 133:2920-2924.
176. Sohail, A., J. Klapacz, M. Samaranayake, A. Ullah, and A. S. Bhagwat. 2003. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Res* 31:2990-2994.
177. Stavnezer, J., and S. P. Bradley. 2002. Does activation-induced deaminase initiate antibody diversification by DNA deamination? *Trends Genet* 18:541-543.
178. Steinel, N. C., and D. I. Bolnick. 2017. Melanomacrophage Centers As a Histological Indicator of Immune Function in Fish and Other Poikilotherms. *Front Immunol* 8:827.
179. Stowers, L., C. Herrnstadt, A. Grothe, E. Pease, M. Osterlund, P. Cable, M. Brolaski, and J. Gautsch. 1992. Rapid isolation of plasmid DNA. *Am Biotechnol Lab* 10:48.
180. Sunyer, J. O. 2013. Fishing for mammalian paradigms in the teleost immune system. *Nat Immunol* 14:320-326.
181. Ta, V. T., H. Nagaoka, N. Catalan, A. Durandy, A. Fischer, K. Imai, S. Nonoyama, J. Tashiro, M. Ikegawa, S. Ito, K. Kinoshita, M. Muramatsu, and T. Honjo. 2003. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat Immunol* 4:843-848.
182. Tarlinton, D. 1998. Germinal centers: form and function. *Curr Opin Immunol* 10:245-251.
183. Tarlinton, D. M., and K. G. Smith. 2000. Dissecting affinity maturation: a model explaining selection of antibody-forming cells and memory B cells in the germinal centre. *Immunol Today* 21:436-441.
184. Tashiro, J., K. Kinoshita, and T. Honjo. 2001. Palindromic but not G-rich sequences are targets of class switch recombination. *Int Immunol* 13:495-505.
185. Taylor, K. L., J. A. Lister, Z. Zeng, H. Ishizaki, C. Anderson, R. N. Kelsh, I. J. Jackson, and E. E. Patton. 2011. Differentiated melanocyte cell division occurs in vivo and is promoted by mutations in Mitf. *Development* 138:3579-3589.
186. Torres, M., and A. Casadevall. 2008. The immunoglobulin constant region contributes to affinity and specificity. *Trends Immunol* 29:91-97.
187. Tort Bardolet, L., J. C. Balasch Alemany, and S. MacKenzie. 2003. Fish immune system. A crossroads between innate an adaptive responses. *Inmunologia* 22:277-286.

188. Vallejo, A. N., C. F. Ellsaesser, N. W. Miller, and L. W. Clem. 1991. Spontaneous development of functionally active long-term monocytelike cell lines from channel catfish. *In Vitro Cell Dev Biol* 27A:279-286.
189. van Ginkel, F. W., N. W. Miller, M. A. Cuchens, and L. W. Clem. 1994. Activation of channel catfish B cells by membrane immunoglobulin cross-linking. *Dev Comp Immunol* 18:97-107.
190. Villota-Herdoiza, D., E. A. Pila, S. Quiniou, G. C. Waldbieser, and B. G. Magor. 2013. Transcriptional regulation of teleost Aicda genes. Part 1 - suppressors of promiscuous promoters. *Fish Shellfish Immunol* 35:1981-1987.
191. Volff, J. N. 2005. Genome evolution and biodiversity in teleost fish. *Heredity (Edinb)* 94:280-294.
192. Wabl, M., M. Cascalho, and C. Steinberg. 1999. Hypermutation in antibody affinity maturation. *Curr Opin Immunol* 11:186-189.
193. Wakae, K., B. G. Magor, H. Saunders, H. Nagaoka, A. Kawamura, K. Kinoshita, T. Honjo, and M. Muramatsu. 2006. Evolution of class switch recombination function in fish activation-induced cytidine deaminase, AID. *Int Immunol* 18:41-47.
194. Wang, C. L., R. A. Harper, and M. Wabl. 2004. Genome-wide somatic hypermutation. *Proc Natl Acad Sci U S A* 101:7352-7356.
195. Wang, C. L., and M. Wabl. 2004. Precise dosage of an endogenous mutagen in the immune system. *Cell Cycle* 3:983-985.
196. Wang, L., W. C. Jackson, P. A. Steinbach, and R. Y. Tsien. 2004. Evolution of new nonantibody proteins via iterative somatic hypermutation. *Proc Natl Acad Sci U S A* 101:16745-16749.
197. Wang, Q., K. R. Kieffer-Kwon, T. Y. Oliveira, C. T. Mayer, K. Yao, J. Pai, Z. Cao, M. Dose, R. Casellas, M. Jankovic, M. C. Nussenzweig, and D. F. Robbiani. 2017. The cell cycle restricts activation-induced cytidine deaminase activity to early G1. *J Exp Med* 214:49-58.
198. Wedemayer, G. J., P. A. Patten, L. H. Wang, P. G. Schultz, and R. C. Stevens. 1997. Structural insights into the evolution of an antibody combining site. *Science* 276:1665-1669.
199. Wilson, M., E. Bengten, N. W. Miller, L. W. Clem, L. Du Pasquier, and G. W. Warr. 1997. A novel chimeric Ig heavy chain from a teleost fish shares similarities to IgD. *Proc Natl Acad Sci U S A* 94:4593-4597.
200. Woo, C. J., A. Martin, and M. D. Scharff. 2003. Induction of somatic hypermutation is associated with modifications in immunoglobulin variable region chromatin. *Immunity* 19:479-489.
201. Yang, F., G. C. Waldbieser, and C. J. Lobb. 2006. The nucleotide targets of somatic mutation and the role of selection in immunoglobulin heavy chains of a teleost fish. *J Immunol* 176:1655-1667.
202. Yang, S. Y., S. D. Fugmann, and D. G. Schatz. 2006. Control of gene conversion and somatic hypermutation by immunoglobulin promoter and enhancer sequences. *J Exp Med* 203:2919-2928.

203. Yeap, L. S., and F. L. Meng. 2019. Cis- and trans-factors affecting AID targeting and mutagenic outcomes in antibody diversification. *Adv Immunol* 141:51-103.
204. Yoshikawa, K., I. M. Okazaki, T. Eto, K. Kinoshita, M. Muramatsu, H. Nagaoka, and T. Honjo. 2002. AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* 296:2033-2036.
205. Yu, K., F. T. Huang, and M. R. Lieber. 2004. DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine. *J Biol Chem* 279:6496-6500.
206. Yu, K., D. Roy, M. Bayramyan, I. S. Haworth, and M. R. Lieber. 2005. Fine-structure analysis of activation-induced deaminase accessibility to class switch region R-loops. *Mol Cell Biol* 25:1730-1736.
207. Zan, H., and P. Casali. 2013. Regulation of Aicda expression and AID activity. *Autoimmunity* 46:83-101.
208. Zarrin, A. A., F. W. Alt, J. Chaudhuri, N. Stokes, D. Kaushal, L. Du Pasquier, and M. Tian. 2004. An evolutionarily conserved target motif for immunoglobulin class-switch recombination. *Nat Immunol* 5:1275-1281.
209. Zarrin, A. A., C. Del Vecchio, E. Tseng, M. Gleason, P. Zarin, M. Tian, and F. W. Alt. 2007. Antibody class switching mediated by yeast endonuclease-generated DNA breaks. *Science* 315:377-381.
210. Zecca, L., D. Tampellini, M. Gerlach, P. Riederer, R. G. Fariello, and D. Sulzer. 2001. Substantia nigra neuromelanin: structure, synthesis, and molecular behaviour. *Mol Pathol* 54:414-418.
211. Zhang, W., P. D. Bardwell, C. J. Woo, V. Poltoratsky, M. D. Scharff, and A. Martin. 2001. Clonal instability of V region hypermutation in the Ramos Burkitt's lymphoma cell line. *Int Immunol* 13:1175-1184.
212. Zhao, Y., Q. Pan-Hammarstrom, Z. Zhao, and L. Hammarstrom. 2005. Identification of the activation-induced cytidine deaminase gene from zebrafish: an evolutionary analysis. *Dev Comp Immunol* 29:61-71.
213. Zhou, H., E. Bengten, N. W. Miller, L. W. Clem, and M. Wilson. 2003. The T cell receptor beta locus of the channel catfish, Ictalurus punctatus, reveals unique features. *J Immunol* 170:2573-2581.
214. Zucca, F. A., E. Basso, F. A. Cupaioli, E. Ferrari, D. Sulzer, L. Casella, and L. Zecca. 2014. Neuromelanin of the human substantia nigra: an update. *Neurotox Res* 25:13-23.
215. Zwollo, P., S. Cole, E. Bromage, and S. Kaattari. 2005. B cell heterogeneity in the teleost kidney: evidence for a maturation gradient from anterior to posterior kidney. *J Immunol* 174:6608-6616.

Appendix

A1.1 Ectopic Expression Of AID In A Channel Catfish B-Cell Line Was Attempted But Curtailed By System Malfunction And Time Constraints

Ectopic expression of AID via a tetracycline-controlled system provides a more reliable and consistent means to increase AID expression in the channel catfish 1B10 immortal B-cell line than does mitogenic stimulation. It also allows for the expression of heterologous and/or modified AID genes for comparative studies within the context of channel catfish B-cells. A Tet-Off system was chosen over a Tet-On system because an experiment by Kuzin et al. [Kuzin, 2001] suggested that the presence of tetracycline or its analogues in the culture medium may damp AID-mediated outcomes. Since it is the outcomes of AID-mediated reactions that were to be examined in this experiment, efforts were made to avoid potentially diminishing what were anticipated to be relatively rare events.

For the initial set of experiments, a mammalian AID (human) and another known teleost AID (zebrafish) were chosen as comparators for channel catfish AID. A carboxy-terminal mutant form of AID (i.e. carboxy-truncated AID) was also included for each species as a means to potentially increase nuclear localization of AID.

Additionally, it was intended that, in subsequent experiments, these full-length and carboxy-truncated AIDs could be used to provide some insight into the evolutionary potential for CSR in the fish system. Even though channel catfish cannot undergo CSR because their IgH genes lack downstream constant regions, they do exhibit some signs that both their cis-acting genetic elements and trans-acting cellular factors might have the potential to contribute to CSR. Specifically, repetitive sequence found in the channel catfish IgH gene J-Cμ intron is vaguely reminiscent of CSR switch regions and the predicted channel catfish AID protein exhibits high conservation with human AID in the carboxy-terminal domain where studies of AID from patients with HIGM2 syndrome and

from artificially generated mutant AID proteins have identified a region that is crucial for CSR but dispensable for SHM [Durandy, 2006; Ta, 2003]. The system should allow for assessment of the CSR competence of the fish AIDs, the channel catfish B-cells, and the channel catfish IgH gene intronic sequences via CSR reporters.

The carboxy-truncated AID sequences were based on an AID mutant (JP8B) identified from a Japanese HIGM2 patient whose aicda gene has a frameshift mutation that results in the replacement of the carboxy-terminal 16 residues of the expressed AID protein with a different sequence of 26 residues [Ta, 2003]. Although its CSR capability is severely compromised (~ 4.5% of wt AID), an *in vitro* assay indicated that the JP8B AID is still able to perform a substantial amount of SHM (~ 70% of wt AID) and Ito et al. found SHM increased significantly over wt AID in an equivalent truncated version, JP8Bdel AID [Ito, 2004; Ta, 2003]. Thus, the carboxy-truncated versions of the various AIDs should sever SHM function from any CSR function. For the purpose of experimental clarity, the JP8B frameshift has been altered to a JP8Bdel truncation that wholly eradicates the terminal 16 residues of the carboxy-terminus [Wakae, 2006]. The JP8Bdel of human AID corresponds to the truncation of the carboxy-terminal 15 amino acids in the fish AID proteins [Wakae, 2006].

As detailed below, the tetracycline-controlled AID-expression system was established in the channel catfish 1B10 immortal B-cell line by sequential integration and selection of the control and response (i.e. expression) plasmids. But there were some unexpected problems and there was insufficient time available to troubleshoot them properly.

Complete integration and expression of the pTet-Off(Blastres) control plasmid were confirmed by PCR, RT-PCR, and sequencing before a clone was chosen for transfection with the pTRE2-zeo-XAID or pTRE2-zeo-XAIDCtrunc response plasmid.

The pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc clones, confirmed by PCR, grown with and without doxycycline (i.e. tetracycline analogue) were examined by RT-PCR for repression and expression, respectively, of their

specific AID. Comparison of the RT-PCR products from the two treatment sets indicated that, in those clones that were able to express their specific AID, the presence of doxycycline in the growth medium did not lead to repression of transcription of AID from the pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc response plasmids. While AID mRNA expression was the same within a clone regardless of whether there was doxycycline in the growth medium, different clones expressed different amounts of AID mRNA as would be expected from random gene integration (e.g. different sites in the genome and variable copy numbers).

The pTRE2-zeo-XAID and pTRE2-zeo-XAIDCtrunc RT-PCR products were purified from the gel and sequenced. Notably, two of the three full-length pTRE2-zeo-CfAID clones had sequence alterations that would result in inactivation of the translated AID protein (i.e. mutations and introduction of a premature stop codon that results in truncation of the carboxy-terminal 63 amino acids in one clone and a 25 nucleotide internal deletion that also caused a frameshift that results in missense after the amino-terminal 64 amino acids in the other clone), but the other AID sequences were accurate. Although the action of catfish AID on the integrated CfAID genes cannot be entirely ruled out here, given that it has previously been noted that transgenes are susceptible to mutation during genome integration [Peters, 1996] and AID-mediated SHM in mammals is not associated with the introduction of relatively large deletions unless the AID is mutated [McBride, 2004], it is likely that the alterations observed in the integrated CfAID genes occurred as a result of the integration process.

Because RT-PCR data indicated that repression was not operational, it is probable that the exogenous AID had been expressed since its integration a few months prior to RT-PCR analysis. As it exceeded the time required to obtain detectable mutations in the PMA/CI stimulation experiment, this should have been sufficient time for SHM to occur in the IgH gene VDJ exon if the AID expression was high enough. But subcloning and sequencing of the endogenous IgH gene VDJ exon from the confirmed full-length CfAID clone did not identify any mutations.

A1.2 Expression System Malfunction Was Not Resolved

The genetic elements of a Tet-Off system for expression of AID from various organisms was established in the channel catfish immortal B-cell line 1B10, which was proven to be SHM competent, to comparatively explore the capacity for AID function within the context of teleost B-cells. A similar Tet-Off system, which abrogates the need to use chemicals that could potentially cloud the AID activity results, has been used successfully in a murine B lymphoma cell line (CH12F3-2) [Muramatsu, 2000]. As both an extension of ongoing research and a complementary companion study to the Wakae et al. [Wakae, 2006] study in mammalian cells, the system was intended to allow for two main lines of AID research within the context of teleost B-cells. First, a comparison of the relative SHM activities of AID molecules and their derivatives from various organisms in the endogenous IgH gene and in SHM reporters within channel catfish B-cells. Additionally, the mutation data generated from the controlled expression of channel catfish AID within these cells would allow for better characterization of AID activity and targeting in a selection-free system [Dancyger, 2012] that could then be more extensively compared with channel catfish *in vivo* SHM data [Yang, 2006] as a means to identify possible selection. Second, assessment, via CSR reporter, of the CSR potential of various AID molecules within channel catfish B-cells. Assuming CSR competence in these cells, the CSR reporter construct could then be altered to incorporate other sequences as switch targets, such as the repetitive sequence from the channel catfish IgH gene J-Cμ intron, to explore the evolution of CSR.

The way in which the Tet-Off system failed to repress the expression of AID in the presence of doxycycline suggests that the problem was likely with the doxycycline. That is, the expression of transgenic AID mRNA is proof that the expression construct (i.e. pTRE2-zeo-XAID or pTRE2-zeo-XAIDCtrunc) is both stably integrated and able to be expressed and also that tTA proteins are present and functional. Thus, the failure in the system is likely due to the only other component, the doxycycline. Although this problem was not resolved due to time

constraints, it is probable that repression could have been achieved with either a higher concentration of doxycycline or the use of a suitable analogue (i.e. tetracycline or anhydrotetracycline). Other reasons for the apparent leakiness (i.e. basal expression) of the expression system cannot be wholly discounted but seem unlikely as no difference in AID expression was observed between the two treatments (i.e. with or without doxycycline) within clones. In theory, unless transcription was at maximum capacity as a result of leakiness, derepression should have led to an observable increase in expression above any transcription due to leakiness as transcription initiated by tTA from the seven tetracycline operator sequences of the Tetracycline Response Element (TRE) is expected to be robust. And it is unlikely that transcription was at maximum capacity in all of the clones as a result of leakiness as different clones expressed different levels of AID mRNA.

That no mutations were identified in the highly transcribed endogenous IgH gene of the pTRE2-zeo-CfAID 1B10 cells, despite being in culture for longer than was required for a readily detectable level of mutations to occur in the PMA/CI-stimulated 1B10 cells or during ectopic expression of AID in higher vertebrate cell lines (i.e. a few days to weeks [Harris, 2002; Martin, 2002; McBride, 2004; Rada, 2002]), may be due to deficient detection (e.g. small sample size sequenced), an incorrect assumption about the level and/or duration of AID expression in the cultures (e.g. doxycycline did initially repress expression but stock degraded over time), or to factors discussed below.

Although the Tet-Off system was chosen deliberately to circumvent potential damping of AID processes by the antibiotic tetracycline [Kuzin, 2001], Bachl et al. [Bachl, 2001] did not find any significant difference in the frequency of mutation when they tested doxycycline in hypermutating 18-81 cells. This suggests that the presence of doxycycline in the media was not the reason for the apparent lack of SHM in this experiment.

If the problem was merely inefficient detection of mutations within an unselected system [Tarlinton, 1998; Zhang, 2001], the introduction of an SHM reporter, an already intended subsequent step, could have overcome this. It

could be that the threshold level of AID expression required for mutation was not met [Michael, 2003; Zhang, 2001], possibly due to inadequate translation or to low copy number of pTRE2-zeo-CfAID integration, although there should be considerable expression driven from the response plasmid even if there is only one copy [Muramatsu, 2000]. Expression analyses relied solely on mRNA, due to the lack of anti-channel catfish AID antibody, and it is possible that the level of channel catfish AID mRNA did not correlate well with the level of AID protein [Dedeoglu, 2004]. Whether for this or other reasons, it is possible that this particular clone was suboptimal for analysis. Clones expressing AID ectopically through a Tet-Off Expression System have been found to vary in their ability to class switch to IgA despite expressing similar levels of germline Cα switch region transcript and apparently similar levels of AID [Muramatsu, 2000]. And Rada et al. [Rada, 2002] found that a majority of G418-resistant AID-GFP fusion transfectants (i.e. selected for stable integration within a site that allows expression from the transgenes) did not produce detectable fluorescence, although they did not investigate this with PCR or RT-PCR.

Although ectopic expression of AID is sufficient to overcome regulatory controls and indiscriminately hypermutate the genome in mammalian cells (see Introduction), it is possible that, in addition to AID expression, channel catfish B-cells need to be in an activated state for SHM to occur. This is another potential explanation, other than levels of constitutive AID expression being below the threshold for SHM, for the lack of mutation in the 1B10 endogenous IgH gene during 10 years in continuous culture. For mammals, it has been postulated that the lack of activating factors (e.g. available antigen, T-cell help, or germinal center environment), in addition to the lack of selection, may contribute to the lower rates of mutation observed in hypermutating cell lines versus *in vivo* SHM [Diaz, 2001; Green, 1998]. Even *in vivo*, SHM occurs at a “lower frequency” in the absence of germinal centers (e.g. in lymphotoxin-alpha-deficient mice) [Diaz, 2001].

If activation of channel catfish B-cells is required for SHM to occur, it could regulate SHM by altering AID directly (e.g. post-translational modification

[McBride, 2006]) or influencing other aspects of the process (e.g. repair). Indeed, it may be that channel catfish B-cells very effectively exclude AID from the nucleus or utilize other means (e.g. proteasomal degradation [Aoufouchi, 2008]) to suppress AID-mediated mutation unless specifically activated. The Tet-Off 1B10 clones expressing the carboxy-terminal truncated AIDs, which lack NES and so are not excluded from the nucleus, were intended, in part, to overcome any potent sequestration strategy used to restrict SHM.

Even though potential causes of the apparent Tet-Off Expression System failure were identified and potential solutions devised, the issues with the system could not be resolved within the time constraints of the project. This meant that the activity of AIDs within the context of a teleost B-cell could not be more fully characterized. An unfortunate setback, but it does not detract from the fact that the data obtained from the 1B10 PMA/CI stimulation experiments was sufficient to confirm that SHM does occur in the endogenous IgH gene in channel catfish B-cells when channel catfish AID is expressed at an adequate level. Furthermore, the data obtained from the channel catfish B-cell line mitogen stimulation experiments contributed to AID expression analyses that facilitated exploration of germinal center evolution, avenues of inquiry that proved to be more productive.

A2 Oligonucleotide Sequences Referenced In The Materials And Methods

The oligonucleotide sequences (i.e. primers) referred to in the Materials and Methods are provided here in a table. The single-column table is divided by bold borders into cells that are subdivided into three rows. Each row contains specific information about an aspect of the oligonucleotide (i.e. name, sequence, or description). A detailed legend for the table is provided in the topmost cell.

Table 3. DNA Oligonucleotides used as primers.

Primer Name The names of the primers are given in the format that was used in the laboratory: P-# accompanied by a brief descriptor that is followed by "+" or "-" symbols to indicate whether the primer sequence is "sense" or "anti-sense", respectively. Primers are listed in numerical order according to P-#.
Primer Sequence The primer sequences are given in the format: 5' Linker-*Endonuclease Site*-**Special Feature**-Gene Sequence.
Primer Description The primer descriptions include information on the gene from which the oligonucleotide primes (e.g. the originating organism or plasmid), salient features regarding its design and use, and its position in the sequence of the given GenBank Accession #.
P-97 Hu AID--
TCA AAG TCC CAA AGT ACG
Human (*Homo sapiens*) AID anti-sense primer. Includes stop codon and the five codons that precede it in the transcript. Position: Base 656 to base 673 in GenBank Accession # AB040431.
P-99 Cc b-actin+
CCA T**C**G AGC ACG GTA TTG TC
Carp (*Cyprinus carpio*) beta-actin sense primer. Also primes from Catfish (*Ictalurus punctatus*) beta-actin. Position: Base 1882 to base 1901 in Carp GenBank Accession # M24113 and base 54 to base 73 in Channel Catfish GenBank Accession # FD308780 with one mispaired base five bases from the 5' end (bold typeface).
P-100 Cc b-actin-
CCA TCT CCT GCT CGA AGT C
Carp (*Cyprinus carpio*) beta-actin anti-sense primer. Also primes from Channel Catfish (*Ictalurus punctatus*) beta-actin. Position: Base 2640 to base 2658 in Carp GenBank Accession # M24113 and base 509 to base 527 in Channel Catfish GenBank Accession # FD308780.

P-124 pCR2.1+
TCC *ACC GGT* CCG AGC GCA GCG AGT CAG TGA G
Invitrogen plasmid pCR2.1 sense primer with 5' 'linker' sequence containing enzyme recognition sequence for AgeI restriction endonuclease. Position: Base 3894 to base 3915 in pCR2.1.
P-134 AID+++
CAA RAA TGT NCG NTG GGC
Composite 'Fish' AID degenerate sense primer to obtain Channel Catfish AID. Designed from alignment of predicted Fugu (*Fugu rubripes*) and Zebrafish (*Danio rerio*) AID transcripts obtained from WGS traces with Human (*Homo sapiens*; GenBank Accession # AB040431), Mouse (*Mus musculus*; GenBank Accession # AF132979), and Chicken (*Gallus gallus*; GenBank Accession # AF059262) AIDs. Position: Degenerate sequences for a portion of the CDS for amino acid sequence YKNVRWA. Base 105 to base 122 (TAAGAATGTGCGCTGGGC) with one mismatch in the Channel Catfish AID mRNA GenBank Accession # AY436507.
P-154 a to l
TTT TTT TTT TTT TTV N
PolyA+ mRNA anti-sense primer (i.e. an oligo dT primer). A mix of equal parts of 12 different primers that comprises all possible combinations of "VN" on the 3' end of the primer for use in the reverse transcription of polyA+ mRNA transcripts.
P-157 fishAID+
GAC TTT GGA CAC CTC GCA ATC GC
Composite 'Fish' AID sense primer to obtain Channel Catfish AID based on the predicted Fugu/Zebrafish AID alignment (see full description with P-134 AID+++). Position: Sequence for a portion of the CDS for amino acid sequence DFGHLRNR. Base 193 to base 215 (GACTTCGGACACCTGCGCAATCG) with three mismatches in the Channel Catfish AID mRNA GenBank Accession # AY436507.
P-160 fishAID--
CAG GGC GAC CAG GAG CAG AAC

Composite ‘Fish’ AID anti-sense primer to obtain Channel Catfish AID based on the Fugu/Zebrafish AID alignment (see full description with P-134 AID+++). Position: Reverse complementary sequence for a portion of the CDS for amino acid sequence WFCSWSPC. Base 344 to base 324 (CAGGGTGACCAGGAACAGAAC) with two mismatches in the Channel Catfish AID mRNA GenBank Accession # AY436507.
P-161 fishAID---
CAG AAT CCG ATT GAG TTT CCG
Composite ‘Fish’ AID anti-sense primer to obtain Channel Catfish AID based on the Fugu/Zebrafish AID alignment (see full description with P-134 AID+++). Position: Reverse complementary sequence for a portion of the CDS for amino acid sequence RKLNRIL. Base 627 to base 607 (CAGGATTCGCTGTAGTTTCCG) with five mismatches in the Channel Catfish AID mRNA GenBank Accession # AY436507.
P-167 CfAID3'RA
CAA GAA ACG CAA CAG TCC CGA CTC GC
Sense primer to obtain Channel Catfish AID in 3’ RACE. Position: Base 159-184 in GenBank Accession # AY436507.
P-170 CfAID5'RACE
GGT CGT CCC AAG CCT TGA AAG CCT TC
Anti-sense primer to obtain Channel Catfish AID in 5’ RACE. Position: Base 580-555 in GenBank Accession # AY436507.
P-183 CfAID++
TCT TTC AGA RTG **ATG** AGC AAG CTG G
Channel Catfish AID sense primer. Includes some of 5' UTR, Kozak sequence, and translation start codon (bold typeface). Position: Base 40 to base 64 in GenBank Accession # AY436507.
P-185 CfAID--
AAT GTT TCA GTG ATT CAA GTC TCA GG

Channel Catfish AID anti-sense primer. From 3' UTR. Position: Base 733 to base 758 in GenBank Accession # AY436507.
P-207 huAID 5'
ACC ACT **ATG** GAC AGC C
Human AID sense primer. Includes some of 5' UTR, Kozak sequence, and translation start codon (bold typeface). Position: Base 71 to base 86 in GenBank Accession # AB040431.
P-219 CMV for
CGC AAA TGG GCG GTA GGC GTG
Cytomegalovirus (CMV) promoter sense primer. Position: Base 623 to base 643 in Clontech plasmid pTet-Off.
P-221 tet +
AGC TGC TTA ATG AGG TCG GAA TC
Tetracycline-responsive transcriptional activator (tTA) sense primer. Position: Base 817 to base 839 in Clontech plasmid pTet-Off.
P-256 pTRE2pur+
CGG TGG GAG GCC TAT ATA AGC AGA G
Clontech plasmid pTRE2pur sense primer. Primer set P-256 and P-257 flank MCS of pTRE2pur for PCR screening and sequencing. Position: Base 330 to base 354 in Clontech plasmid pTRE2pur.
P-257 pTRE2pur-
GAA AAC TTT GCC CCC TCC ATA TAA C
Clontech plasmid pTRE2pur anti-sense primer. Primer set P-256 and P-257 flank MCS of pTRE2pur for PCR screening and sequencing. Position: Base 607 to base 631 in pTRE2pur.
P-263 CfVH1pr++

GGT ATA WAC AGC ATG TTY TTG
Channel Catfish Immunoglobulin Heavy Chain VH1 family promoter degenerate sense primer. In the rearranged IgH gene of 1B10, this primer sequence is 33-13bp 5' of the transcription start site and 103-83bp 5' of the translation start codon. Position: Base 294 to base 314 in GenBank Accession # AF068137.
P-268 3'ofCfJH1-
TAT AGG AAT AAG AAT AGG AAC AG
3' of Channel Catfish Immunoglobulin Heavy Chain Joining Region 1 anti-sense primer. Position: Base 1530 to base 1552 in GenBank Accession # L22340.
P-271 pTRE2pur+
CGA ATT CGA GCT CGG TAC
Clontech plasmid pTRE2pur sense primer. For use in PCR screening to discriminate between vector-generated transcripts and those from endogenous genes. Position: Base 449 to base 466 in pTRE2pur.
P-283 5'UTRcfAID
TTT GCA TGA CTC AAC CAC TG
Channel Catfish AID sense primer. Position: Base 9 to base 28 in GenBank Accession # AY436507.
P-313 tet--
CTG GCT CTG CAC CTT GGT GAT C
Tetracycline-responsive transcriptional activator (tTA) anti-sense primer. Position: Base 1305 to base 1326 in Clontech plasmid pTet-Off.
P-335 5'tetTAHindIII+
AGA TCT T*AA GCT* *T*GT TGA CAT TGA TTA TTG ACT AG
Clontech plasmid pTet-Off sense primer with 5' 'linker' sequence containing enzyme recognition sequence for HindIII restriction endonuclease. Primer set P-335 and P-336 is

to amplify tetTA with its associated promoter and polyA signal from pTet-Off. Position: Base 85 to base 107 in pTet-Off.
P-336 3'tetTANheI-
TAT GCG GCC *GCT AGC TAG C*TT GGT CGA GCT GAT ACT TC
Clontech plasmid pTet-Off anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequence for NheI restriction endonuclease. Primer set P-335 and P-336 is to amplify tetTA with its associated promoter and polyA signal from pTet-Off. Position: Base 2232 to base 2253 in pTet-Off.
P-391 5'zfAID++
CGA CTT TCG GA**A TG**A TCT GCA AGC TGG ACA GTG
Zebrafish AID sense primer. Includes some of 5' UTR, Kozak sequence, and translational start codon (bold typeface). Position: Base 34 to base 66 in GenBank Accession # AY528720.
P-392 3'zfAID---
CAT CAT AAC CCA AGA AGA GCA AAA ACA TCC CTC AGA TC
Zebrafish AID anti-sense primer. Position: Base 642 to base 679 in GenBank Accession # AY528720.
P-393 CfCu1BaNo+
ATA T*GG ATC CGC GGC CG*C GGT TTA GTC ACT CTT GGC TGC GTC AC
Channel Catfish Immunoglobulin Heavy Chain μ Constant Exon 1 sense primer with 5' 'linker' sequence containing enzyme recognition sequences for BamHI and NotI restriction endonucleases. Position: Base 520 to base 546 in GenBank Accession # M27230.
P-394 CfCu3RsXb-
ATA *TTC TAG ACG GTC CG*G TGC CAT TGA TCC ACT CTT CGA AAC
Channel Catfish Immunoglobulin Heavy Chain μ Constant Exon 3 anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequences for XbaI and RsrII

restriction endonucleases. Position: Base 1278 to base 1304 in GenBank Accession # M27230.
P-395 CfAIDBaNo+
ATA T*GG ATC CGC GGC CGC* ATG AGC AAG CTG GAC AGT GTG CTG CTG AC
Channel Catfish AID sense primer with 5' 'linker' sequence containing enzyme recognition sequences for BamHI and NotI restriction endonucleases. Position: Base 52 to base 80 in GenBank Accession # AY436507.
P-396 CfAIDRsXb-
ATA T*TC TAG ACG GTC CG*T TAA AGG CCC AGC AGA GCG AAG CCA TC
Channel Catfish AID anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequences for XbaI and RsrII restriction endonucleases. Position: Base 655 to base 681 in GenBank Accession # AY436507.
P-429 BaNo5'cfAID+
TTT T*GG ATC CGC GGC CGC* TCT TTC AGA RTG **ATG** AGC AAG CTG G
Channel Catfish AID sense primer with 5' 'linker' sequence containing enzyme recognition sequences for BamHI and NotI restriction endonucleases. This is modified P-183 which includes some of 5' UTR, Kozak sequence, and translational start codon (bold typeface). Position: Base 40 to base 64 in GenBank Accession # AY436507.
P-430 EcHi3'cfAIDCtrnc-
TTT T*GA TAT CAA GCT T***TT A***AG GCT GCA GGA TTC GCT GTA G
Channel Catfish AID anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequences for EcoRV and HindIII restriction endonucleases. This primer inserts a premature stop codon (bold typeface) into the Channel Catfish AID gene. This leads to the production of Channel Catfish AID protein with a C-terminal deletion of 15 amino acids (JP8Bdel truncation). Position: Base 613 to base 633 in GenBank Accession # AY436507.
P-431 BaNo5'zfAID+

TTT T*GG ATC CGC GGC CGC* CGA CTT TCG GA**A TG**A TCT GCA AGC TGG ACA GTG
Zebrafish AID sense primer with 5' 'linker' sequence containing enzyme recognition sequences for BamHI and NotI restriction endonucleases. This is modified P-391 which includes some of 5' UTR, Kozak sequence, and translational start codon (bold typeface). Position: Base 34 to base 66 in GenBank Accession # AY528720.
P-432 EcHi3'zfAIDCtrnc-
TTT T*GA TAT CAA GCT T***TT A**AG GCT GCA GAA TCC GAT TGA G
Zebrafish AID anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequences for EcoRV and HindIII restriction endonucleases. This primer inserts a premature stop codon (bold typeface) into the Zebrafish AID gene. This leads to the production of Zebrafish AID protein with a C-terminal deletion of 15 amino acids (JP8Bdel truncation). Position: Base 609 to base 630 in GenBank Accession # AY528720.
P-433 BaNo5'HuAID+
TTT T*GG ATC CGC GGC CGC* ACC ACT **ATG** GAC AGC CTC TTG ATG
Human AID sense primer with 5' 'linker' sequence containing enzyme recognition sequences for BamHI and NotI restriction endonucleases. This is modified P-207 huAID which includes some of 5' UTR, Kozak sequence, and translational start codon (bold typeface). Position: Base 73 to base 97 in GenBank Accession # NM_020661.
P-434 EcHi3'HuAIDCtrnc-
TTT T*GA TAT CAA GCT T***TT A**GG GCA AAA GGA TGC GCC GAA GCT G
Human AID anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequences for EcoRV and HindIII restriction endonucleases. This primer inserts a premature stop codon (bold typeface) into the Human AID gene. This leads to the production of Human AID protein with a C-terminal deletion of 16 amino acids (JP8Bdel truncation). Position: Base 602 to base 625 in GenBank Accession # NM_020661.
P-460 XhSazeo+

TTT T*CT CGA GTC GAC* AAT TCT GTG GAA TGT GTG TCA GTT AG
Zeocin resistance gene sense primer with 5' 'linker' sequence containing enzyme recognition sequences for XhoI and SalI restriction endonucleases. Position: Base 1794 to base 1819 in Invitrogen plasmid pcDNA4/TO.
P-461 XhSazeo-
TTT T*CT CGA GTC GAC* GGT ATA CAG ACA TGA TAA GAT AC
Zeocin resistance gene anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequences for XhoI and SalI restriction endonucleases. Position: Base 2896 to base 2868 in Invitrogen plasmid pcDNA4/TO.
P-466 pCR2.1delreg-
GCC AGT GTG ATG GAT ATC TGC
Invitrogen plasmid pCR2.1-TOPO sense primer in region found by Forbes et al. [Forbes, 2007] to be frequently excised in aberrant clones. Position: Base 308 to base 328 in pCR2.1-TOPO.
P-467 SpXbcfTCRalpha+
TAT A*AC TAG TCT AGA* CAG AAG AGA AAC GGG AGC CGT CAA TTT AC
Channel Catfish T-cell Receptor Alpha sense primer with 5' 'linker' sequence containing enzyme recognition sequences for SpeI and XbaI restriction endonucleases. The PCR product of P-467 with either P-468 or P-469 span two introns. Position: Base 392 to base 420 in GenBank Accession # U58505.
P-468 SpXbcfTCRalpha-
TAT A*AC TAG TCT AGA* CTA GCT CAT CCA GAC TTT CAG CGT CAT C
Channel Catfish T-cell Receptor Alpha anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequences for SpeI and XbaI restriction endonucleases. The PCR product of P-467 and P-468 spans two introns. Position: Base 717 to base 745 in GenBank Accession # U58505.
P-469 SpXbcfTCRa3'UTR-

TAT A*AC TAG TCT AGA* CAT GTT GTC ATG TTA TGC AGA AC
Channel Catfish T-cell Receptor Alpha anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequences for SpeI and XbaI restriction endonucleases. The PCR product of P-467 and P-469 spans two introns. Position: Base 823 to base 845 in GenBank Accession # U58505.
P-472 EcpGEM3zf+
TA*G ATA TC*G TTT TCC CAG TCA CGA CGT TG
Promega plasmid pGEM-3zf(+) sense primer with 5' 'linker' sequence containing enzyme recognition sequence for EcoRV restriction endonuclease. Position: Base 3138 to base 3158 in pGEM-3zf(+).
P-473 EcpGEM3zf-
TA*G ATA TC*C TAT GAC CAT GAT TAC GCC AAG
Promega plasmid pGEM-3zf(+) anti-sense primer with 5' 'linker' sequence containing enzyme recognition sequence for EcoRV restriction endonuclease. Position: Base 89 to base 110 in pGEM-3zf(+).
P-484 pGEM3zfT7pro
CGG CCA GTG AAT TG**T AAT ACG ACT CAC TAT AGG**
Promega plasmid pGEM-3zf sense primer. This primer includes most of the T7 phage RNA polymerase promoter sequence of the plasmid (bold typeface). Position: Base 3167 to base 2 in pGEM-3zf(+).
P-485 pGEM3zfSP6pro
CGC CAA GCT **ATT TAG GTG ACA CTA TAG AA**T ACT C
Promega plasmid pGEM-3zf anti-sense primer. This primer includes the SP6 phage RNA polymerase promoter sequence of the plasmid (bold typeface). Position: Base 62 to base 95 in pGEM-3zf(+).
P-564 pTRE2zeoseq+
TGT TTT GAC CTC CAT AGA AGA CAC

Clontech plasmid pTRE2pur-derived plasmid pTRE2zeo sense primer. Position: Base 399 to base 422 in pTRE2pur.
P-565 pTRE2seq-
CTC CAT ATA ACA TGA ATT TTA CAA TAG
Clontech plasmid pTRE2pur-derived plasmid pTRE2zeo sense primer. Position: Base 591 to base 617 in pTRE2pur.
P-566 T7cfTCRbeta+
TAA TAC GAC TCA CTA TAG GGA GAG GAC TCT ATA ACA GGA CCA AAA C
Channel Catfish T-cell Receptor Beta sense primer for both alleles of TCRbeta 1 with 5' 'linker' sequence containing the T7 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-610 with 'T7 linker'. Position: Base 713 to base 736 in allele 1 GenBank Accession # U39193 and base 683 to base 706 in allele 2 GenBank Accession # U58508.
P-567 SP6cfTCRbeta-
ATT TAG GTG ACA CTA TAG AAG AGA ATA TAG CAC TCA TAT CTG AAC TAG
Channel Catfish T-cell Receptor Beta anti-sense primer for both alleles of TCRbeta 1 with 5' 'linker' sequence containing the SP6 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-611 with 'SP6 linker'. Position: Base 13275 to base 13299 in allele 2 GenBank Accession # AF410785. Due to high sequence identity between the two alleles [Zhou, 2003] and the truncation of the available 3' UTR sequence of allele 1 in the NCBI database, this primer was designed from only allele 2 sequence.
P-568 SP6cfTCRbeta--
ATT TAG GTG ACA CTA TAG AAG AGC CAA ATT CCT CAC TGT AAT GAA G
Channel Catfish T-cell Receptor Beta anti-sense primer for both alleles of TCRbeta 1 with 5' 'linker' sequence containing the SP6 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-612 with 'SP6 linker'. Position: Base 13504 to base 13526 in allele 2 GenBank Accession # AF410785. Due to high sequence identity between the two alleles [Zhou, 2003] and the truncation of the

available 3' UTR sequence of allele 1 in the NCBI database, this primer was designed from only allele 2 sequence.
P-569 T7IpCD4L1+
TAA TAC GAC TCA CTA TAG GGA GAG TGG CAC ACA CAC CTT GCT CAT G
Channel Catfish CD4-Like 1 sense primer with 5' 'linker' sequence containing the T7 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-613 with 'T7 linker'. Position: Base 817 to base 840 in GenBank Accession # DQ435301.
P-570 SP6IpCD4L1-
ATT TAG GTG ACA CTA TAG AAG AGT CAG GTC TTG TAG AAA CCT TTG
Channel Catfish CD4-Like 1 anti-sense primer with 5' 'linker' sequence containing the SP6 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-614 with 'SP6 linker'. Position: Base 1395 to base 1416 in GenBank Accession # DQ435301.
P-571 T7IpCD4L2+
TAA TAC GAC TCA CTA TAG GGA GAC CAT CAG GAA CAG TGC GCA GAG
Channel Catfish CD4-Like 2 sense primer with 5' 'linker' sequence containing the T7 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-615 with 'T7 linker'. Position: Base 175 to base 196 in GenBank Accession # DQ435302.
P-572 SP6IpCD4L2-
ATT TAG GTG ACA CTA TAG AAG AGC ATG AGT CTT AAC TTC CCT TG
Channel Catfish CD4-Like 2 anti-sense primer with 5' 'linker' sequence containing the SP6 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-616 with 'SP6 linker'. Position: Base 575 to base 595 in GenBank Accession # DQ435302.
P-573 T7cfAID5'+
TAA TAC GAC TCA CTA TAG GGA GAC TCA GTT TTG CAT GAC TCA AC

Channel Catfish AID sense primer with 5' 'linker' sequence containing the T7 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-617 with 'T7 linker'. Position: Base 1 to base 23 in GenBank Accession # AY436507.
P-574 SP6cfAIDN-
ATT TAG GTG ACA CTA TAG AAG AGG TTG GGC ATC TGA GAC ATG AAG
Channel Catfish AID sense anti-primer with 5' 'linker' sequence containing the SP6 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-618 with 'SP6 linker'. Position: Base 372 to base 395 in GenBank Accession # AY436507.
P-575 T7cfAIDC+
TAA TAC GAC TCA CTA TAG GGA GAG GTG TGC AAG TGA CAG TCA TGA C
Channel Catfish AID sense primer with 5' 'linker' sequence containing the T7 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-619 with 'T7 linker'. Position: Base 483 to base 506 in GenBank Accession # AY436507.
P-576 SP6cfAID3'-
ATT TAG GTG ACA CTA TAG AAG AGC ACC ATA ATT GAT TGT CAT G
Channel Catfish AID anti-sense primer with 5' 'linker' sequence containing the SP6 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-620 with 'SP6 linker'. Position: Base 697 to base 718 in GenBank Accession # AY436507.
P-577 T7RAG1+
TAA TAC GAC TCA CTA TAG GGA GAG CTG TTT CGA GTT CGT TCA TTG
Channel Catfish Recombination-Activating Gene 1 sense primer with 5' 'linker' sequence containing the T7 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-621 with 'T7 linker'. The RAG1 genes of some fish, including Catfish, differ from those found in other organisms in that they have an intron; the PCR product of P-577 and P-578 spans the intron.

Position: Base 5 to base 27 in GenBank Accession # AY423858.
P-578 SP6RAG1-
ATT TAG GTG ACA CTA TAG AAG AGA TGG TCT GTC TCC ACA CTC TAT TG
Channel Catfish Recombination-Activating Gene 1 anti-sense primer with 5' 'linker' sequence containing the SP6 phage RNA polymerase promoter site (bold typeface) recommended by Ambion. This is P-622 with 'SP6 linker'. The RAG1 genes of some fish, including Channel Catfish, differ from those found in other organisms in that they have an intron; the PCR product of P-577 and P-578 spans the intron. Position: Base 630 to base 653 in GenBank Accession # AY423858.
P-579 P472T7
GTT TTC CCA GTC ACG ACG TTG TAA TAC GAC TCA CTA TAG GGA GA
T7 phage RNA polymerase promoter site (recommended by Ambion) sense primer with 5' 'linker' sequence containing the basic P-472 primer sequence. P-579 and P-580 are for PCR amplification of either PCR products or plasmid inserts generated with primers containing T7 and SP6 sites. The P-579/P-580 PCR products can then be sub-cloned into pCR2.1-TOPO and PCR amplified out with P472/473. This helps to avoid issues with the T7 site when it occurs in both the vector and the insert.
P-580 P473SP6
CTA TGA CCA TGA TTA CGC CAA G ATT TAG GTG ACA CTA TAG AAG AG
SP6 phage RNA polymerase promoter site (recommended by Ambion) anti-sense primer with 5' 'linker' sequence containing the basic P-473 primer sequence. P-579 and P-580 are for PCR amplification of either PCR products or plasmid inserts generated with primers containing T7 and SP6 sites. The P-579/P-580 PCR products can then be sub-cloned into pCR2.1-TOPO and PCR amplified out with P472/473. This helps to avoid issues with the T7 site when it occurs in both the vector and the insert.
P-610 cfTCRbeta+
GGA CTC TAT AAC AGG ACC AAA AC
Channel Catfish T-cell Receptor Beta sense primer for both alleles of TCRbeta 1. This is P-566 without 'T7 linker'.

Position: Base 714 to base 736 in allele 1 GenBank Accession # U39193 and base 684 to base 706 in allele 2 GenBank Accession # U58508.
P-611 cfTCRbeta-
AAT ATA GCA CTC ATA TCT GAA CTA G
Channel Catfish T-cell Receptor Beta anti-sense primer for both alleles of TCRbeta 1. This is P-567 without 'SP6 linker'. Position: Base 13275 to base 13299 in allele 2 GenBank Accession # AF410785. Due to high sequence identity between the two alleles [Zhou, 2003] and the truncation of the available 3' UTR sequence of allele 1 in the NCBI database, this primer was designed from only allele 2 sequence.
P-612 cfTCRbeta--
CCA AAT TCC TCA CTG TAA TGA AG
Channel Catfish T-cell Receptor Beta anti-sense primer for both alleles of TCRbeta 1. This is P-568 without 'SP6 linker'. Position: Base 13504 to base 13526 in allele 2 GenBank Accession # AF410785. Due to high sequence identity between the two alleles [Zhou, 2003] and the truncation of the available 3' UTR sequence of allele 1 in the NCBI database, this primer was designed from only allele 2 sequence.
P-613 IpCD4L1+
GTG GCA CAC ACA CCT TGC TCA TG
Channel Catfish CD4-Like 1 sense primer. This is P-569 without 'T7 linker'. Position: Base 818 to base 840 in GenBank Accession # DQ435301.
P-614 IpCD4L1-
TCA GGT CTT GTA GAA ACC TTT G
Channel Catfish CD4-Like 1 anti-sense primer. This is P-570 without 'SP6 linker'. Position: Base 1395 to base 1416 in GenBank Accession # DQ435301.
P-615 IpCD4L2+
CCA TCA GGA ACA GTG CGC AGA G

Channel Catfish CD4-Like 2 sense primer. This is P-571 without ‘T7 linker’. Position: Base 175 to base 196 in GenBank Accession # DQ435302.
P-616 IpCD4L2-
CAT GAG TCT TAA CTT CCC TTG
Channel Catfish CD4-Like 2 anti-sense primer. This is P-572 without ‘SP6 linker’. Position: Base 575 to base 595 in GenBank Accession # DQ435302.
P-617 cfAID5'+
CTC AGT TTT GCA TGA CTC AAC
Channel Catfish AID sense primer. This is P-573 without ‘T7 linker’. Position: Base 3 to base 23 in GenBank Accession # AY436507.
P-618 cfAIDN-
GTT GGG CAT CTG AGA CAT GAA G
Channel Catfish AID anti-sense primer. This is P-574 without ‘SP6 linker’. Position: Base 372 to base 393 in GenBank Accession # AY436507.
P-619 cfAIDC+
GGT GTG CAA GTG ACA GTC ATG AC
Channel Catfish AID sense primer. This is P-575 without ‘T7 linker’. Position: Base 484 to base 506 in GenBank Accession # AY436507.
P-620 cfAID3'-
CAC CAT AAT TGA TTG TCA TG
Channel Catfish AID anti-sense primer. This is P-576 without ‘SP6 linker’. Position: Base 697 to base 716 in GenBank Accession # AY436507.
P-621 RAG1+
GCT GTT TCG AGT TCG TTC ATT G
Channel Catfish Recombination-Activating Gene 1 sense primer. This is P-577 without ‘T7 linker’. The RAG1 genes of some fish, including Channel Catfish, differ from those

found in other organisms in that they have an intron; the PCR product of P-621 and P-622 spans the intron. Position: Base 6 to base 27 in GenBank Accession # AY423858.
P-622 RAG1-
ATG GTC TGT CTC CAC ACT CTA TTG
Channel Catfish Recombination-Activating Gene 1 anti-sense primer. This is P-578 without 'SP6 linker'. The RAG1 genes of some fish, including Channel Catfish, differ from those found in other organisms in that they have an intron; the PCR product of P-621 and P-622 spans the intron. Position: Base 630 to base 653 in GenBank Accession # AY423858.