The mechanisms of thermosensitivity of human and bony fish activation-induced cytidine deaminase (AID)

by

© Mussa Farag Ali Suliman (**B.Sc**.)

A thesis submitted to the School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Science in Medicine

Immunology and Infectious Diseases Research Program

Division of BioMedical Sciences

Faculty of Medicine

Memorial University of Newfoundland

May 2014

St. John's Newfoundland and Labrador Canada

Abstract

Activation-induced cytidine deaminase (AID) is an enzyme that initiates the secondary antibody diversification events of somatic hypermutation (SHM) and class switch recombination (CSR). AID is expressed in activated B lymphocytes and deaminates cytidines to uridines, thus introducing mutations and/or double-strand breaks throughout the immunoglobulin (Ig) locus. Although AID activity is mostly targeted to the Ig locus, it mutates many other genes at lower rates and this off-target activity induces genomic lesions leading to transformation of B cells. The structure of AID has not been solved, which is partially the reason that several important questions remain unanswered about its mechanism of action. Our lab has recently shown that AID from different species has disparate rates of deamination, and that the deamination rate of AID from each species is thermosensitive, such that fish AID are cold-adapted relative to human AID. Here, we have measured the activity of purified AID from human, zebrafish, channel catfish, pufferfish (takifugu), and hybrid AID containing various parts of the aforementioned, over a wide range and fine increments of temperatures. Our aim was to elucidate the relationship between thermosensitivity of AID and its primary structure, by identifying region(s) and/or residue(s) within the enzyme that affect thermosensitivity. We found that this characteristic is mostly determined by the C-terminus in the sense that swapping this region resulted in transferring thermosensitivity patterns from the donor enzyme. We also found that AID activity is not always optimal at native physiological temperature, such that human AID is optimal at 32°C instead of 37°C. Based on our data, we postulated that this finding may have implications for evolution of AID. Overall, our work has shed light on mechanistic features of AID as well as aspects on its evolution.

Acknowledgements

I would like to thank my supervisor, Dr. Mani Larijani. It has been a great privilege for me to spend two years of my academic career in his lab. He patiently provided the vision, encouragement and advice necessary for me throughout my time in the program.

I would like also to thank my committee members, Dr. Michael Grant, and Dr. Rodney Russell for their time, advice, and helpful suggestions.

Extended gratitude is deserved of my fellow lab mates in the Larijani Lab for their support and technical assistance.

I would like also to acknowledge the advice and guidance of my friends, who have been an essential part of my life in Canada.

Last but not least, I would like to express my appreciation to my family, they deserve my sincerest gratitude. A special thank goes to every one of them, their moral support and assistance has meant more to me than I could ever express.

Table of Contents

List of figures

List of Abbreviations and Symbols

µg	Microgram
µl	Microlitre
^{o}C	Degrees Celsius
A	Adenine
A3F	Apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3F
A3G	Apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3G
Ab	Antibody
AID	Activation-induced cytidine deaminase
Amp	Ampicillin
APOBEC	Apolipoprotein B mRNA-editing, enzyme-catalytic polypeptide-like
BCL-6	B-cell lymphoma 6
BCR	B-cell receptor
BER	Base excision repair
BL	Burkitt's lymphoma
Bp	Base pair
C	Cytidine
C region	Constant region of Ig
CLL	Chronic lymphocytic leukemia
CSR	Class switch recombination
Dr-AID	Zebrafish AID

DLCL	Diffuse large B cell lymphoma
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
DSB	Double stranded break
DTT	Dithiotheritol
EDTA	Ethylene diamine tetra acetic acid
FL	Follicular lymphoma
g	Gram
G	Guanine
GST	Glutathione *S*-transferase
Hs-AID	Human AID
Ig	Immunoglobulin
Ip-AID	Catfish AID
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Lysogeny broth (Luria-Bertani medium)
mL	Millilitre
MMR	Mismatch repair
mRNA	Messenger ribonucleic acid
NES	Nuclear export signal
NLS	Nuclear localization signal
nt	Nucleotide
OD	Optical density
PBS	Phosphate buffered saline

PCR	Polymerase chain reaction
RAG-1	Recombination activating gene product 1
RAG-2	Recombination activating gene product 2
RNA	Ribonucleic acid
S	Switch region of Ig
ssDNA	Single stranded DNA
SDS	Sodium dodecyl sulfate
SDS PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SHM	Somatic hypermutation
SOC	Super Optimal Broth
T	Thymidine
tadA	tRNA adenosine deaminase
TBE	Tris borate/EDTA electrophoresis buffer
Tc	Cytotoxic T cell
TCR	T cell receptor
Th	T helper cell
U	Uridine
UDG	Uracil-N-DNA glycosylase
V	Variable region of Ig
V(D)J	Recombination of variable (V), diversity (D), and joining (J) gene segments
WRC	A hot spot motif targeted by AID where W=A/T, R= A/G, and C= cytidine
Zn	Zinc

1. Introduction

1.1 The immune system: innate and adaptive immunity

The immune system is comprised of a collection of biological structures and processes that function together to provide protection against a wide variety of insults; such as viruses, bacteria, tumour cells, and parasites (Warrington et al., 2011; Lanasa & Weinberg, 2011). There are two arms: The innate immune system and the adaptive immune system. The innate immune system is the first line of defense against foreign pathogens. This system is comprised of antigen-independent defense mechanisms such as chemical and physical barriers (e.g. lysozymes and mucous membranes), plasma proteins (complement), and cells (e.g. natural killer cells, macrophages, and dendritic cells). These mechanisms respond immediately or within a short time to the invading pathogens in a non-specific way, but do not last long and have no immunologic memory. In contrast, adaptive immune responses are the second line of defense and take days or weeks to develop, but last long and have memory to recognize subsequent encounters with the same antigen and respond in a stronger and faster manner. There are two types of adaptive immunity characterized as antigen-specific responses: First, humoral immune responses, mediated by B lymphocytes, second, cell-mediated immune responses, mediated by T lymphocytes (Warrington et al., 2011; Litman et al., 2005; Fleisher & Tomar, 1997).

B and T lymphocytes are the main cellular components of the adaptive immune system. T lymphocytes mature in the thymus and have several subsets: T helper cells (T_h cells) and cytotoxic T cells (also known as T_C or Cytotoxic T Lymphocytes, CTLs)

are the major subtypes of T lymphocytes. In response to antigens T helper cells produce a wide range of cytokines that induce, promote and direct immune responses, whereas cytotoxic T cells produce effective and potent cytotoxins that target and kill infected cells. On the other hand, B lymphocytes mature in the bone marrow, and leave it after maturation. Naive B cells, which have never been exposed to an antigen, express unique membrane-bound antibodies (antigen binding receptors) on their surfaces. These primary antigen binding receptors have low binding affinity to specific foreign pathogens. Once these cells are activated by a specific antigen, they migrate to germinal center structures in lymph nodes. There, they proliferate and differentiate into antibody-producing B cells (plasma cells) that produce large amounts of antibodies, which function to neutralize the activating antigen(s). However, some of the B cells do not differentiate into antibody-producing plasma cells; instead they differentiate into long-lived memory cells, which become the effector cells of subsequent immunological memory. Besides producing antibodies, B cells also produce and release cytokines and function as antigen presenting cells (Warrington et al., 2011; Litman et al., 2005; Fleisher & Tomar, 1997; Gellert, 2002).

There are five subclasses of antibodies (Immunoglobulins) known as IgM, IgA, IgG, IgD, and IgE. Immunoglobulins are glycoprotein molecules each monomer of which is comprised of four polypeptide chains: Two identical heavy chains and two identical light chains linked together by disulphide bonds to form a "Y" shaped protein (figure 1). The amino acid sequences at the end of the heavy and light chains vary extensively among immunoglobulins to form a antigen-binding site that confers to the antibody its unique antigen binding properties. The bottom of the "Y" shaped molecule, known as Fc

region (constant region) determines the effector function of the immunoglobulin by interacting with different cell surface receptors. The Fc portion is identical within antibodies of the same isotype and varies only from one isotype to another (Woof & Burton, 2004; Fleisher & Tomar, 1997).

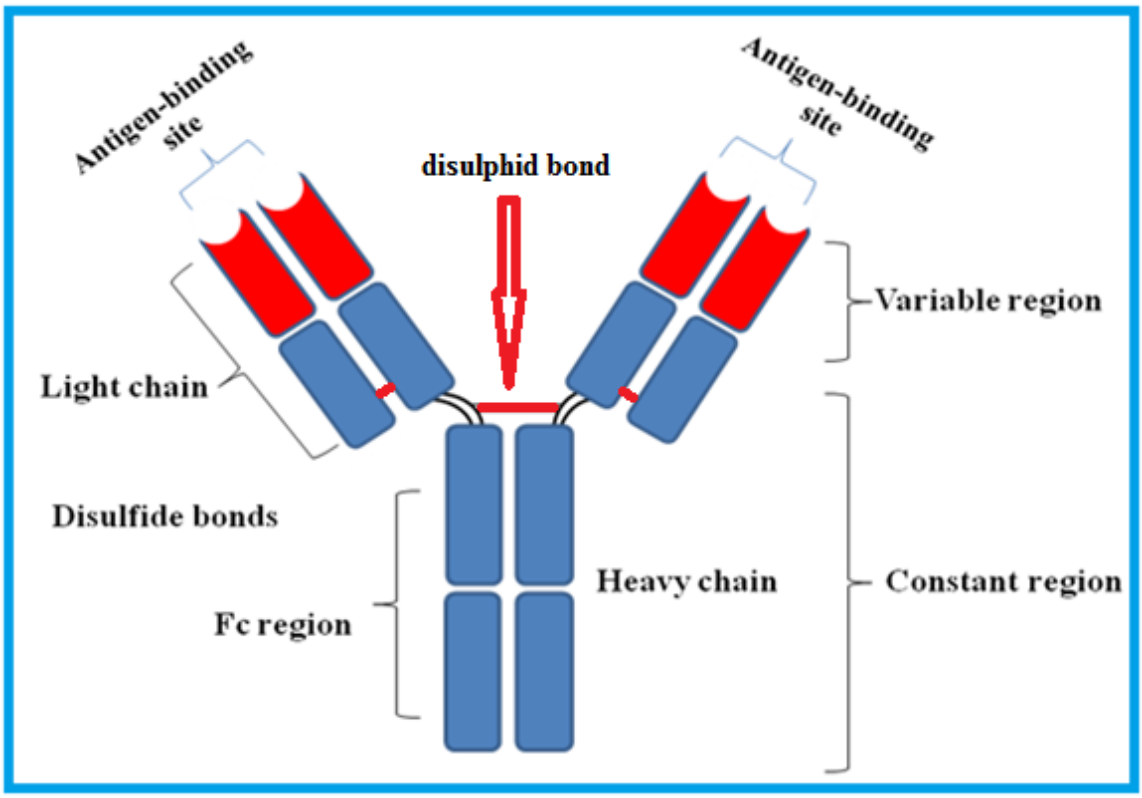


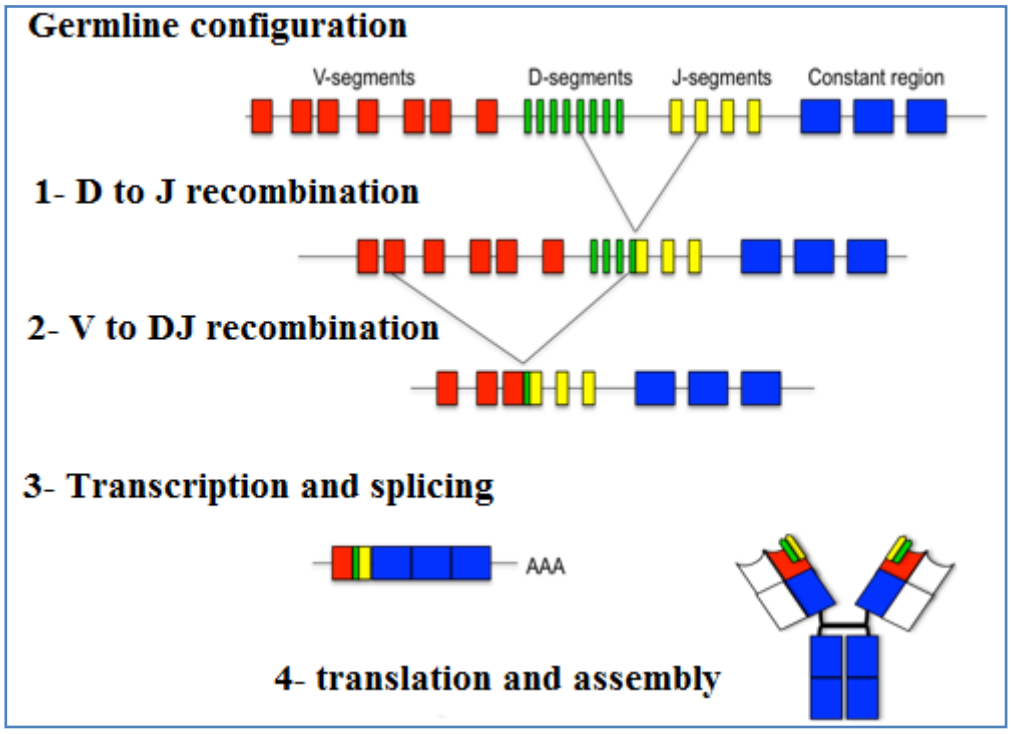
Figure 1: Schematic diagram of IgG.

Figure 1: Schematic diagram of IgG.

Each immunoglobulin (Ig) monomer consists of four polypeptides: two identical heavy chains and two identical light chains joined to form a "Y" shaped molecule. The variable regions of light and heavy chains form the antigen-binding sites of the Ig molecules. Each Ig contains two antigen-binding sites. The Fc region plays a significant role in determining the effector functions of immunoglobulin.

1.2 V(D)J recombination: primary antibody diversification

The vertebrate immune system generates great diversity in its pre-immune antigen receptors even before encountering antigens. This massive repertoire of antigen-binding receptors is achieved through a process known as V(D)J recombination (V=variable, D=diversity, and J= joining gene segments). V(D)J rearrangement is a highly specialized and site-specific DNA recombination process that occurs during early B and T lymphocyte development to create functional immunoglobulins (Igs) and T cell receptors (TCRs). This genetic rearrangement is mediated by two lymphoid-specific recombination activating gene products known as RAG-1 and RAG-2. This sophisticated mechanism is the only recognized site-specific rearrangement process in higher eukaryotes. In mammals, there are seven well-identified antigen receptors loci: T cell receptor α, β, γ, and δ loci, as well as κ, λ, and IgH loci. During B and T lymphocyte development, random recombination of V, D and J gene segments joins them together at the DNA level to be transcribed and then translated to form the variable and the constant regions of the B and T cell receptors (figure 2). In addition to this combinatorial diversification mechanism, there is considerable imprecision in the joining of V, D and J segments that further adds diversity to regions of the gene that encode antigen binding sites of B and T cell receptors (Gellert, 2002; Matsuda et al., 1998).



(Modified from Janway 2001)

Figure 2: V(D)J recombination.

Figure 2: V(D)J recombination.

V(D)J recombination is a type of site-specific DNA rearrangement that takes place in primary lymphoid organs (bone marrow and thymus) early during B and T cell development. V(D)J recombination brings together gene segments to encode functional immunoglobulins (Igs) and T cell receptors (TCRs) in B and T cells, respectively. During B cell maturation, gene segments for both heavy (VDJ) and light (VJ) chains are randomly assembled. Each B cell ends up with functional genes that code for light and heavy chains. Heavy and light chains associate to form an antibody as a membrane bound receptor. The cleavage phase of V(D)J recombination is mediated by two recombination activating gene products known as RAG-1 and RAG-2. The joining phase is mediated by non homologous end joining (NHEJ)-type DNA repair processes that are not cell type-specific.

1.3 Secondary diversification of the antibody repertoire

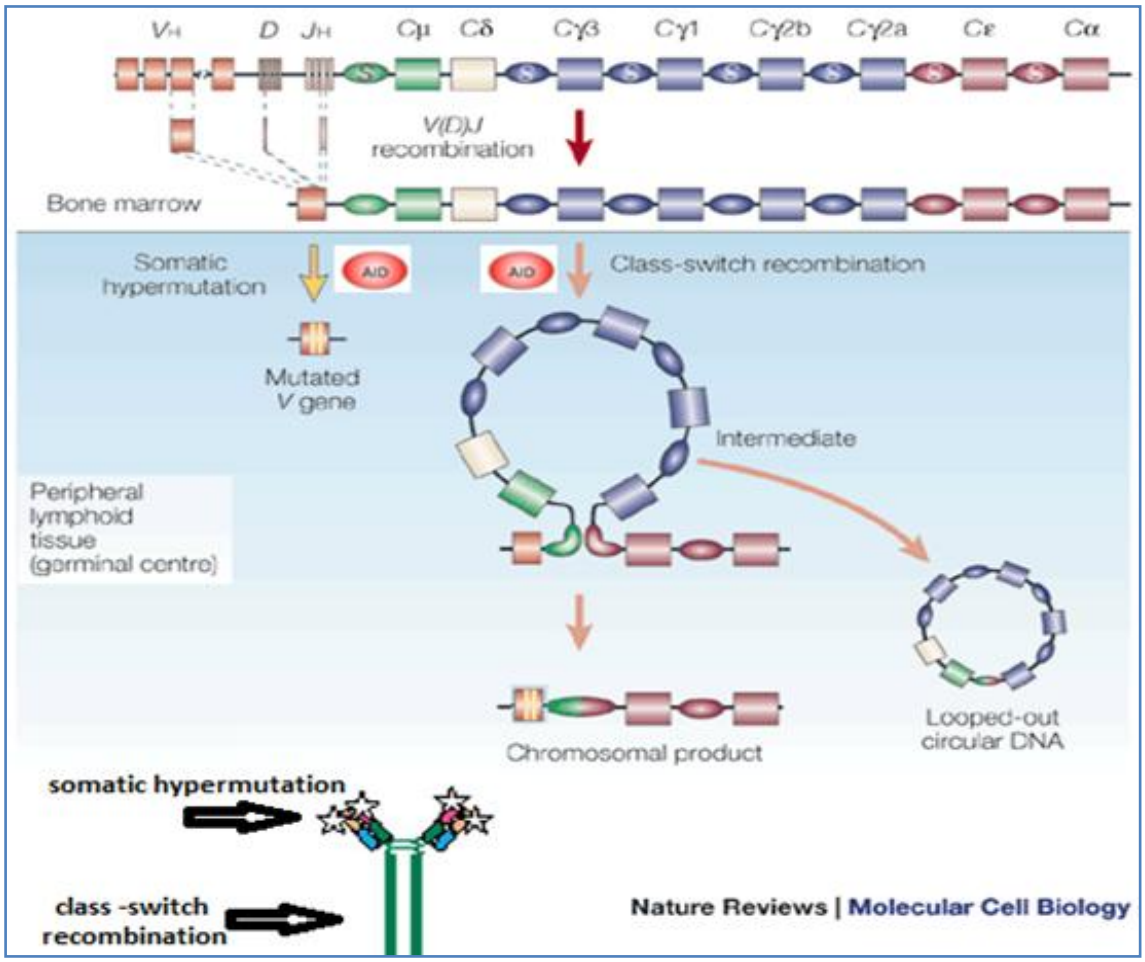
V(D)J recombination generates diversity in the primary antibody repertoire of developing B lymphocytes in the bone marrow. These pre-immune B cell receptors on the surface of the naive B cells are produced without previous encounter with any antigens. Despite the fact that the antigen binding sites of these initial receptors cover a wide spectrum of specificities, they have low affinities for any specific antigen. However, following exposure to an immunogen, B cells get activated, and secondary antibody diversifications take place to generate antibodies with switched isotype and increased binding affinity for specific antigenic determinants. This further diversification occurs when activated B cells migrate to the germinal centers of the lymph nodes, and the antibody genes are further mutated and/or recombined through processes known as somatic hypermutation (SHM) and class switch recombination (CSR) (figure 3) (Muramatsu et al., 2000; Okazaki et al., 2002; Revy et al., 2000).

SHM is the mechanism by which point mutations are introduced in the rearranged variable V(D)J exon of immunoglobulin genes. The rate of these random mutations in activated B cells is approximately one million times higher than the rate of spontaneous mutations in somatic cells (Odegard & DG, 2006; Rajewsky, 1996). These point mutations will result in some B cell receptors with improved affinity and others with reduced affinity for a specific antigen. Due to enhanced binding to a limited pool of antigen, B cells that express antigen receptors on their membrane with high affinity will be positively selected while other B cells expressing receptors with lower affinity to antigens will be eliminated by apoptosis. The processes of secondary antibody diversification, followed by cellular selection for B cells that express antibodies with

enhanced affinity for antigens, are collectively known as affinity maturation (Or-Guil et al., 2007; Rajewsky, 1996; Neuberger et al., 2000; Ziqiang et al., 2004). Besides affinity maturation, activated B cells will also undergo CSR. This mechanism involves recombination of the constant (*C*) region of the antibody genes with downstream *C* regions, thus replacing the primary *C* region (*C*μ heavy-chain). This results in a switch from the initial IgM isotype to IgA, IgE, or IgG. Since the antigen binding site does not change during CSR, the antigen specificity of the antibody will not be altered. Instead, the antibody will be able to interact with various effector molecules to carry out different effector functions (Stavenzer et al., 2008; de Yebenes & Ramiro, 2006).

The secondary antibody diversification events of SHM and CSR are initiated by an enzyme called activation-induced cytidine deaminase (AID) (Muramatsu et al., 1999; Muramatsu et al., 2000; Revy et al., 2000; Martin et al., 2002; Okazaki et al., 2002; Barreto et al., 2005; Neuberger et al., 2005). AID is expressed specifically in activated B lymphocytes; the processes of SHM and CSR takes place only in these cells. The significance of AID in triggering the secondary antibody diversification was revealed by interrupting the expression of AID in mice, which resulted in lack of class switch recombination and somatic hypermutation. In addition, individuals with inherited defects in the *AID* gene also lack CSR and SHM, which results in an accumulation of IgM and absence of affinity maturation. Individuals with such a disorder suffer from a syndrome known as hyper IgM type2 immunodeficiency**.** Generally, in hyper IgM syndrome, B lymphocytes continue producing IgM antibodies, leading to high levels of IgM, and a lack of all other antibody isotypes (IgA, IgG, and IgE). This results in recurrent bacterial infections, and other diseases as well as lymph node hyperplasia caused by giant germinal

centers, presumably because B cells do not undergo proper affinity maturation, or because AID activity can also result in B cell death, either through mutating its antibody so that it no longer recognizes antigen, or through its off-target mutating genotoxic effects in B cells (Durandy et al., 2006; Dooley et al., 2006; Revy et al., 2000).



(Modified from Kinoshita & Honjo, 2001)

Figure 3: Somatic hypermutation (SHM) and class switch recombination (CSR).

Figure 3: Somatic hypermutation (SHM) and class switch recombination (CSR).

The Secondary antibody diversification events of somatic hypermutation (SHM) and class switch recombination (CSR) are initiated by an enzyme called activation-induced cytidine deaminase (AID). These processes occur in the peripheral lymphoid tissues (germinal centers), whereas V(D)J recombination occurs in the bone marrow. SHM introduces point mutations in the rearranged variable *V(D)J* exon of immunoglobulin genes, leading to a tremendous number of variants of the recombined antibodies some of which change the initially weak and/or broadly binding antibodies into stronger and more specifically-binding variants. CSR involves double strand breaks (DSB) in switch (*S*) sequences upstream of constant (*C*) exons, followed by recombination. CSR brings the downstream constant (*C*) region in the proximity of the *V* region by mediating recombination between *Sμ* and another downstream *S* region. Deleted DNA is looped out as circular DNA. This process results in the antibody isotype being changed from IgM to IgA, IgE, or IgG.

1.4 The mechanisms of the mutagenic activity of AID in initiating secondary antibody diversification and oncogenesis

Ig genes are the main targets of AID. It plays a vital role in achieving antibody diversity by initiating the DNA lesions that get SHM and CSR underway (Conticello et al., 2007; Okazaki et al., 2003)**.** Several lines of evidence showed that AID can bind and target single-stranded DNA (ssDNA), but not double-stranded DNA (dsDNA) or RNAs (Larijani et al., 2007a; Dickerson et al., 2003; Bransteitter et al., 2003). Previous work in the field showed that cytidines within trinucleotide WRC (W=A or T, R= A or G) motif are preferentially mutated by AID *in vitro* (Beale et al., 2004; Pham et al., 2003). This preferential mutating by the purified AID *in vitro* is consistent with the pattern of SHM *in vivo* (Larijani et al., 2005; Yang & Schatz, 2007). The ssDNA specificity of AID along with its preference for highly transcribed genes has led to the idea that transcription is a major source of ssDNA substrate for AID activity. It is thought that during the transcription of Ig genes, the DNA double helix of the Ig genes become temporarily unwound and single-stranded locally, generating ssDNA that AID can bind and mutate, causing the formation of a uracil-guanine mismatch (U:G) mismatch in the target DNA (Martin & Scharff, 2002; Li et al., 2002). This U:G mismatch can be recognized and processed by various DNA repair mechanisms such as mismatch repair (MMR) or base-excision repair (BER) pathways (David et al., 2007; Fischer et al 2007; Guo-Min, 2008). These types of DNA repair processes are often error prone because they recruit error-prone polymerases to fill in the gaps. As a result, the initial cytidine to uridine mutations mediated by AID can manifest in the full spectrum of point mutations (i.e. from any

nucleotide to any of the other three) at variable region genes (Li et al., 2004; Peled et al., 2008; Neuberger et al., 2005).

Because of genomic instabilities induced by the mutagenic activity of AID, its expression must be firmly controlled under physiological conditions (Gonda et al., 2003; Okazaki et al., 2003; Babbage et al., 2006). However, recent studies indicated that AID is responsible for genomic instability, chromosomal translocations, and the development of some cancers (Robbiani et al., 2008; Mechtcheriakova et al., 2012; Feldhahn et al., 2007; Takai, et al., 2012). Although the mutagenic activities of AID are mostly directed to Ig genes, AID can also act on non-Ig genes (Pasqualucci et al., 2001; Liu et al., 2008; Colotta et al., 2009). This off-target activity of AID can cause mutations and chromosomal translocations that lead to B cell leukemias and lymphomas, including Burkitt's lymphoma which is caused by recombination between the proto-oncogene *c-Myc* and the *IgH* locus S region (Ramiro et al., 2006; Robbiani et al., 2008). This translocation puts the expression of the proto-oncogene under the control of the potent *IgH* enhancers, and results in constitutive and/or over expression of this gene. Mouse model studies have shown that AID can directly cause DSBs at the Ig locus and c-Myc breakpoints and that in the absence of AID, this translocation is not detectable in B cells (Ramiro et al., 2006; Dorsett et al., 2007; KuÈppers & Dalla-Favera, 2001). Another example of lymphoma-associated chromosomal translocations mediated by AID is the *BCL-2-IgH* translocation observed in follicular lymphomas. Also, the double stranded breaks initiated by AID in *BCL6* and *IRF4* loci can result in chromosomal translocations leading to transformation of normal mature B cells to diffuse large B cell lymphoma (DLBCL) and multiple myeloma (Greeve et al., 2003; Jankovic et al., 2010; Park, 2012;

Pasqualucci et al., 2001). In addition, recent studies have revealed that AID may be indelibly expressed in cells other than B lymphocytes, such as T cells and non-immune cells, where it may also mediate tumorigenesis (Park 2012; Komori et al., 2008). Factors such as microbial infections and pro-inflammatory cytokines in non-immune cells have been shown to induce AID expression (Okazaki et al., 2007; Komori et al., 2008; Grivennikov et al., 2010).

1.5 The AID/APOBECs family of cytidine deaminase enzymes

AID belongs to the APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) family of cytidine deaminase enzymes. AID/APOBEC enzymes have in common the capability to deaminate cytidine to uridine, thereby introducing mutations within DNA or RNA. Members of this family in humans include: APOBEC1, APOBEC2, APOBEC3A, APOBEC3C, APOBEC3D, APOBEC3F, APOBEC3G, APOBEC3H, APOBEC4, and AID (Conticello et al., 2007; Conticello, 2008; Larijani & Martin, 2012). The first characterized family member was APOBEC1, an enzyme responsible for converting cytidine at position 6666 of the apolipoprotein B (ApoB) pre-mRNA to uridine, thus generating a premature stop condon and resulting in a truncated ApoB protein (Navaratnam et al.,1993a; Teng et al., 1993). It has been shown that both APOBEC3G and APOBEC3F play a crucial role in innate antiretroviral immunity by deaminating cytidines in the minus-sense-strand of retroviral genomic DNA, therefore restricting replication of these viruses (Chiu & Greene, 2008; Sheehy et al., 2002; Yu et al., 2004). An important characteristic of APOBEC3G, APOBEC3F, and AID is their preference to target the third cytidine within specific trinucleotide motifs (Sheehy et al., 2002; Yu et al., 2004; Beale et al., 2004; Conticello et al., 2005; Conticello et al., 2007; Larijani & Martin, 2012). The specific sequences in which cytidine is preferentially targeted and deaminated by A3G, A3F, and AID comprise CCC, TTC, WRC (W= A or T, R= A or G), respectively (Pham et al., 2003; Larijani et al., 2005; Sohail et al., 2005; Conticello et al., 2005). The preferential targeting of the cytidine located within these particular motifs by A3G, A3F, and AID enzymes is an inherent feature that correlates qutie closely with the sequence specificity of these enzymes *in vivo* (Larijani et al., 2005;

Coker & Petersen, 2007; Pham et al., 2003). The exact physiological functions of other APOBEC family members such as APOBEC2 and APOBEC4 have not yet been elucidated (Liao et al., 1999; Anant et al., 2001). Previous studies on the AID/APOBEC family using phylogenetic sequence analysis have revealed that the ancestral family members are AID and APOBEC2 since these are the only members to date identified in the more evolutionarily distant cartilaginous and bony fish (Conticello, 2008; Conticello et al., 2005; Barreto & Magor, 2011; Pancer & Cooper, 2006; Larijani & Martin, 2012).

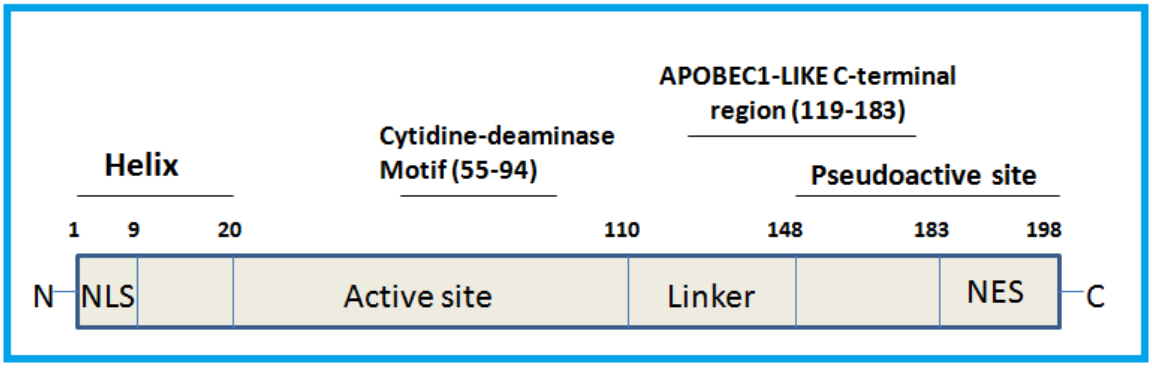
1. 6 Structural features and biochemical parameters of human AID

Purification of an active and pure AID enzyme is a difficult procedure, largely because it is a highly positively charged molecule that forms non-specific associations with other proteins and also because its expression is inherently highly genotoxic to cells. For these reasons, the crystal structure of AID has not been solved yet. In the absence of an experimentally determined structure, various structural features of AID may be modeled based on comparing the primary and secondary structure of this enzyme with other solved APOBEC family members and other deaminases such as the bacterial tRNA adenosine deaminase (tadA) whose structures have been solved (Losey et al., 2006; Prochnow et al., 2007; Conticello, 2008; Larijani & Martin; 2012; Dancyger et al., 2012; Chen et al., 2008; Holden et al., 2008). Several putative regions of AID have been suggested, which are responsible for different functions (figure 4). Among the functional regions identified thus far are an APOBEC-like region and nuclear export signal (NES), Zn-coordinating region, and the N terminus with a putative nuclear localization signal (NLS) (Barreto & Magor, 2011; Larijani & Martin, 2012). The putative NLS as well as regions suggested to be important for dimerization and tetramerization are located within the N terminal half of AID, while the NES is located in the C-terminus region. Recent studies suggested that AID can shuttle between the cytoplasm and nucleus using its NLS, NES, and other hydrophobic amino acids of the C-terminal region (Patenaude et al., 2009; Ito et al., 2004; Barreto & Magor, 2011). The function of the APOBEC-like region has not yet been characterized. This region is located at positions 119-183 of the carboxyl terminal end, and was defined based on comparing the sequence alignment of AID to that of other APOBECs (Barreto & Magor, 2011). The putative catalytic region of AID has

also been identified. It is the Zn-coordinating region of the enzyme ranging from residue 56 to 90 and responsible for the enzymatic activity (Conticello, 2008; Barreto & Magor, 2011). This region contains the canonical Zn-coordinating H[A/V]E-x[24–36]–PCxxC motif. The two Cysteines (C) and the Histidine (H) within this motif coordinate a zinc cation to form the catalytic core of AID (Figure 5) (Conticello, 2008; Barreto & Magor, 2011). The cytidine binds in this pocket (Larijani & Martin, 2012) and gets deaminated and converted to a uridine through a nucleophilic attack by an activated water molecule (coordinated via the zinc atom) on the ammonium group of the fourth carbon atom of the cytidine, while the nearby glutamate functions as a proton donor (Conticello, 2008; Abdouni et al., 2013). Also, two putative DNA binding grooves have been identified that pass over the catalytic pocket of AID (Dancyger et al., 2012; Larijani & Martin, 2012) each of them is about 6 Å in width. Since dsDNA is about 20 Å wide, the size of these grooves is large enough to accommodate only ssDNA but not dsDNA (Pham et al., 2003; Sohail et al., 2003; Larijani et al., 2007a). In addition, the high percentage of positively charged residues, such as Arginine (R) and Lysine (K), at the putative DNA-binding grooves, may further explain why AID has high affinity binding to negatively charged ssDNA (Larijani et al., 2007a).

Previous work by our group has demonstrated that human AID has a number of unique enzymatic characteristics such as a very slow rate of deamination activity, requiring several minutes to perform 1 reaction (most other enzymes perform thousands of reactions/sec) (Larijani, et al., 2007). It has been suggested that the slow deamination rate of human AID and the unusual high affinity binding of ssDNA (nM range affinity for ssDNA) may be an intrinsic property of this enzyme to restrict and control its off-target

mutagenic activity *in vivo* and explain its apparent enzymatic processivity *in vitro* (Pham et al., 2007; Coker & Petersen-Mahrt, 2007; Larijani & Martin, 2012; Dancyger et al., 2012)

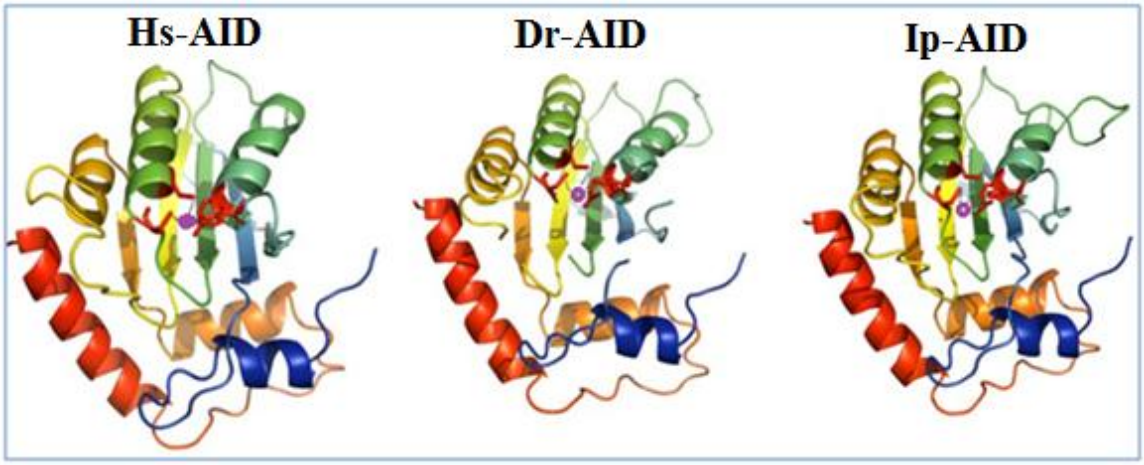


(Modified from Chaudhuri & Frederick, 2004)

Figure 4: A simplified schematic representation of the primary structure of AID.

Figure 4: A simplified schematic representation of the primary structure of AID.

The primary structure of activation-induced cytidine deaminase (AID) is shown, depicting the nuclear-localization sequence (NLS) in the N terminus, nuclear-export sequence (NES) in the C terminus and the catalytic (Zn-coordinating) region between the N and the C terminal regions. The APOBEC1-like C terminal region is located at positions 119-183 of the carboxyl terminal region.



(Modified from Dancyger et al., 2012)

Figure 5: Ribbon models of predicted structures of Hs-AID, Dr-AID, and Ip-AID.

Figure 5: Ribbon models of predicted structures of Hs-AID, Dr-AID, and Ip-AID.

Ribbon models of the predicted monomer structures of human AID (Hs-AID), zebrafish AID (Ip-AID), and catfish AID (Ip-AID) were created based on the solved structure of the APOBEC3G catalytic region. N- to C-terminal progression is shown in color from blue to red. The blue-colored α-helix structure represents the N-terminal region, whereas the red-colored α-helix structure represents the C-terminal region. Putative catalytic (Zn-coordinating) residues are shown in stick format in red. Putative position of the Zn cation is shown by a purple ball.

1.7 Biochemical features of AID orthologs compared to human AID

AID is present in organisms as far back on the evolutionary scale as the earliest gnathostomes (jawed vertebrates) studied to date (Rogozin et al., 2007; Saunders et al., 2010). Even though the 3-dimentional structure of AID is not known, the primary sequence of AID from mammals, amphibians, and birds is highly conserved implying that many structure-function features may also be shared. AID sequence has been identified from many different species including, human, mouse, rat, dog, cat, rabbit, duck, cow and chimpanzee (Barreto & Magor, 2011; Bascove & Frippiat, 2010; Zhao et al., 2005). The oldest living vertebrates for which AID structure-function relationships have been studied are bony fish. Since they are so distantly related, bony fish AIDs are the most divergent from their human counterpart (Conticello et al., 2005; Wakae et al., 2006; Barreto & Magor, 2011) and recent work from our lab has shown that they have altered biochemical characteristics (Dancyger et al., 2012; Abdouni et al., 2013).

It has been demonstrated that AID from various species have disparate deamination rates, and that the deamination rate of AID from each species is thermosensitive (temperature-dependent) and optimal at natural physiological temperatures, such that fish AID are cold-adapted (Dancyger et al., 2012). In general, the cold-adapted catalytic activity of fish AID is ascribed to increased ssDNA binding affinity at lower temperatures. In addition, zebrafish AID has a higher catalytic rate than human AID while catfish AID has the lowest overall level of activity. This discrepancy is mediated by a single amino acid difference in the C terminus (Dancyger et al., 2012).

1.8 Mechanisms of enzyme thermosensitivity

Overall enzyme structure and function are significantly influenced by temperature. For proper enzyme function, a balance has to be maintained between flexibility and stability, particularly in regions which are required for catalysis. Changes in temperature tremendously interrupt this balance. Therefore, enzymes are often temperature-dependent in substrate binding and catalysis. Generally, as temperature increases over a specific range, an enzyme becomes highly flexible; disrupting the catalytic site and other regions which can possibly also diminish the enzyme-substrate binding affinity. If temperature is further increased, the enzyme will eventually denature. On other hand, the enzyme can be catalytically inactive and excessively rigid at cold temperatures, which lowers the catalytic efficiency of the enzyme beneath the level needed to maintain metabolic homeostasis within the cell (Jaenicke 1991; Fields et al., 2001; Fields, 2002; Zavodszky et al., 1998; Field & Houseman, 2004).

Since different organisms have evolved in dissimilar thermal environments, their enzymes may have adapted to disparate temperature conditions. Therefore, enzyme thermosensitivity usually varies among unrelated species (Feller & Gerday, 2003; Smalas, 2000; D'Amico et al., 2002; Somero, 2004; Siddiqui & Cavicchioli, 2006). Also, it has been shown that cold-adapted enzymes in general have distinct properties from warm-adapted ones. On one hand, cold-adapted enzymes have evolved a variety of characteristic structural features that provide them with high flexibility at low temperatures. The high level of plasticity is mainly around the catalytic site of the cold-adapted enzyme, which is usually associated with weak substrate binding affinity and

low-activation energy, but high specific activity at low temperatures (Siddiqui & Cavicchioli, 2006; Fields & Somero, 1998; D'Amico et al., 2002). On the other hand, warm-adapted enzymes (thermostable enzymes) are characterized by reduced activity and low structural flexibility at low temperature (Eijsink et al., 2004; Siddiqui & Cavicchioli, 2006; Fields; D'Amico et al., 2002). Therefore, investigating the thermosensitivity characteristics of a given enzyme in detail can be a useful approach to gain more information about its structure.

1.9 Purpose of study: mapping the thermosensitivity of AID and characterizing mechanisms of thermosensitivity

Rationale:

Our lab has recently shown that human AID is more enzymatically active at 37°C, and has diminished enzymatic efficacy at low temperatures such as 25°C & 18°C. On the other hand, fish AIDs (e.g. zebrafish and channel catfish AID) are catalytically optimal at 25°C, and exhibit lower activity levels at 37°C (Dancyger et al., 2012). Also, a previous study by Conticello and his group on non-mammalian AID has compared the enzymatic activity of the pufferfish (cold-adapted) AID to that of mouse (warm-adapted) AID using a bacterial rifampicin-rescue mutation assays (Conticello, et al., 2005). Their study revealed that at 37°C pufferfish AID has no noticeable deamination activity, but at 18°C this enzyme showed increased deamination activity level (Conticello et al., 2005; Barreto & Magor, 2011). In general, bony fish are categorized as poikilotherms (cold-blooded), organisms whose internal body temperature varies according to the temperature of their surroundings. In contrast, mammals are homeothermic (warm-blooded) organisms that maintain a constant internal body temperature regardless of environmental temperatures. Several earlier studies have shown that various species (e.g. cold- and warm-adapted organisms) that have evolved in different thermal environments have enzymes that are adapted to the distinct temperatures (Feller & Gerday, 2003; Smalas, 2000; D'Amico et al., 2002; Somero, 2004; Siddiqui & Cavicchioli, 2006). Generally speaking, there are significant differences between cold- and warm-adapted

enzymes in their overall structural flexibility, thermostability and substrate binding at certain temperatures (Feller & Gerday, 2003; Georlette et al., 2004).

The molecular basis of AID's cold- and warm-adaption has not yet been elucidated. In this study, we focused our research on better understanding AID thermosensitivity among various species. In particular, we conducted a comparative study to elucidate this characteristic of human AID versus the most divergent bony fish AIDs.

1.10 Specific aims and hypothesis

The specific goals of our research included:

1) To precisely determine the optimal temperature and the thermosensitivity patterns for each AID enzyme.

2) To identify and highlight the important structural region(s) responsible for thermosensitivity of AID.

3) To quantitatively understand thermosensitivity by characterizing the degree of difference in enzyme velocity when measured at optimal and sub-optimal temperatures.

To scrutinize the mechanism underlying thermosensitivity of AID and gain more insights into its structure-function relationships, we used a bacterial expression system to express and purify human, zebrafish, channel catfish, and pufferfish AIDs, along with hybrid AID enzymes made of different portions of the above mentioned enzymes. These hybrid AIDs were used to pinpoint and narrow down regions that can transfer this property of AID. All purified AID enzymes were examined over a wide range of temperatures to precisely determine the optimal temperatures and then thermosensitivity patterns of each AID enzyme. One possible model is that the thermosensitivity of AID is

a characteristic that is based in one continuous region of the enzyme and can be transferred from one species to another by swapping said region. Alternatively, this property might be attributed to the overall conformational structure of AID in a way that requires residues that are discontinuous in the primary structure, to function in combination. Our results suggest that the thermosensitivity of AID is mainly a C-terminal region based feature, and swapping the N-terminal or the catalytic region of AID between species does not influence the thermosensitivity. Importantly, we also found that the optimal temperature for human AID *in vitro* is at 32°C rather than 37°C, which suggests that optimal temperature can be distinct form physiological temperature and provides clues on the evolution of AID.

2. Materials and Methods

2.1 AID expression and purification of AID

The open reading frames (flanked by EcoRI recognition sequences) encoding zebrafish AID (Dr-AID), catfish AID (Ip-AID), and pufferfish AID (Tr-AID) were cloned into pGEX-5x-3 N-terminal GST-tag to create GST-AID as previously described (Dancyger et al., 2012; Larijani et al., 2007; Saunders & Magor, 2004; Zhao et al., 2005; Ichikawa et al., 2006). We utilized EditSeq Software to design hybrid AID enzymes involving different AID regions from human and either zebrafish or channel catfish or pufferfish (takifugu) AIDs. Open reading frames for these hybrid AIDs were designed and synthesized (Genscript) and cloned into pGEX-5-3-GST expression vector by using EcoRI restriction sites in the same way as the original cloning for human AID (GST-Hs AID), zebrafish AID (GST-Dr AID) and catfish AID (GST-Ip AID) (Dancyger et al., 2012; Larijani et al., 2007; Abdouni et al., 2013). Constructs containing AID inserts were verified by restriction digests and sequencing. 1 μg of each GST-AID plasmid DNA was mixed with 50 μL of BL21 (DE3), chemically competent *E. coli*. The reaction was placed for 30 minutes on ice, then heat shocked for 90 seconds at 42°C in preheated water bath, and then 250 μL of Super Optimal Broth (SOB medium) (2% Bacto-tryptone, 0.5% Bacto-yeast extract, 0.05% Sodium chloride, 2.5 mM KCl, and 10 mM MgCl2) was quickly added to each reaction tube. All the reaction tubes were incubated in a shaking incubator at 37°C for 1 hour. 100 μL of each tube was plated on an LB plate with antibiotic (37g LB, agar, 1 litter of MilliQ dH2O, 100 μg/ml ampicillin). All LB plates were inverted and incubated for 12-16 hours at 37°C. Glycerol stocks of bacteria were

frozen for future AID production and purification (glycerol stocks were maintained at -80°C). AID expression was induced and GST-AID was purified using GST columns as previously described (Dancyger et al., 2012; Larijani et al., 2007). Briefly, one colony from each plate was picked (or cultures were started from frozen glycerol stocks) and grown in 250 ml LB medium plus 50 mg/ml Ampicillin at 37°C, 225 rpm until log phase was reached. 250 μL of [1 M] IPTG (Isopropyl β-D-1-thiogalactopyranoside) as well as 250 μL of 50 mg/mL Ampicillin were added to induce AID expression in DE3 bacterial cells. Bacterial cultures were then incubated for 16 hours at 16°C. The next day, each culture was poured into 250 mL centrifuge bottles, and centrifuged for 10 minutes at 5000 rpm, 4°C. The supernatant was drained and pellets were resuspended in 20 mL of cold 1X PBS, then the resuspended pellets were transferred to labelled 50 mL Falcon tubes. Cells were lysed twice in a French pressure cell press (Thermospectronic), then transferred to 50 mL centrifuge tubes and centrifuged at 4°C, 5000 rpm for 10 minutes. After centrifugation, the supernatant was applied to a column of Glutathione-Sepharose beads (Amersham) as per the manufacturer's instructions. The flow through was reapplied to the column, then the column was washed with 1X PBS. AID was eluted with elution buffer (50mM Tris pH 8.0, 100 mM reduced glutathione) as per the manufacturer's recommendations. 15-20 fractions of 500 μL of eluted GST-AID were aliquoted into 1.5 ml tubes kept on ice. The optical density of each eluted fraction was measured using a NanoDrop to determine protein content. AID fractions with the highest concentrations which constitute the peak protein containing fractions were pooled and placed into sealed SnakeSkin® Pleated dialysis tubing and dialyzed overnight in 1 litter of dialysis Buffer

(20 mM Tris-Cl pH 7.5, 100 mM NaCl, 1 mM dithiotheritol (DTT). On the following day, the dialysis buffer was replaced, and the AID fractions were further dialyzed for 3-5 hours. Aliquots of 30-50 μL of the dialysed AID were distributed into 1.5 ml tubes, flash frozen in liquid nitrogen and stored at -80°C for future use in enzyme assays. For each GST-AID enzymes, multiple independent purifications (2-4) were made. The purity of each purified GST-AID protein was assessed by SDS-PAGE. We then investigated the optimal temperature of each hybrid AID in parallel with the purified wild type AID from human and bony fish, over a wide range and fine increments of temperatures.

2.2 AID substrate: preparation and labelling

Generation of a bubble substrate as a target for AID deamination has been previously described (Larijani & Martin, 2007). Briefly, this substrate is 56 nt in length and comprises a 7 nt long single-stranded bubble region that has cytidine (C) within the hotspot motif WRC (WRC motifs W=A/T, R=purine (A/G), C=cytidine). The top strand of the substrate (BHagtop) was purified by desalting, whereas the bottom strand containing the hot spot motif WRC (TGCbub7, AGCbub7, TACbub7) was purified by High Performance Liquid Chromatography (HPLC).

The top strand (BHagtop)

5'–AGA TCC TGC CCC GGC ACT TCG CCC GGG TTT TTC CAG TCC CTT CCC GCT TCA GTG AC– 3'

The bottom strand **TGC**bub7

5'–GTC ACT GAA GCG GGA AGG GAC TGT G**TGC***TT CCG GGC GAA GTG CCG GGG CAG GAT CT– 3'

The bottom strand **TAC**bub7

5’–GTC ACT GAA GCG GGA AGG GAC TGT G**TAC***TT CCG GGC GAA GTG CCG GGG CAG GAT CT– 3’

The bottom strand **AGC**bub7

5’–GTC ACT GAA GCG GGA AGG GAC TGT G**AGC***TT CCG GGC GAA GTG CCG GGG CAG GAT CT– 3’

Each 1 µL of 2.5 pmol of the bottom strand (target strand) was labelled with 3 µL of 15 µCi/µL Adenosine 5′-triphosphate, ATP-[γ-32P] at the 5′-end using 1 µL of polynucleotide kinase (PNK) (NEB, Ipswich, MA, USA), 4 µL of autoclaved MilliQ dH2O, and 1 µL of 10XPNK buffer. The reaction was mixed well and incubated for 1 hour at 37°C, then at 65°C for 10 minutes to deactivate the PNK enzyme. The reaction was briefly centrifuged, 10 µL of 1X Tris-EDTA (TE) buffer was mixed with the reaction, and then the 20 µL reaction was added to a mini-Quick spin DNA column (Roche) and spun at 3000 rpm for 3 minutes to separate and remove any unlabelled substrate. 5 µL 1 M KCl, 3 µL 2.5 pmol/µL BHagtop top strand (complementary strand) were added to the resulting flow through to form the bubble substrate. The final volume was adjusted to 50 µL by adding autoclaved MilliQ water. This mixture was incubated for 2 minutes at 96 °C, and then the temperature of the incubation was decreased by 1°C every 30 seconds until the temperature reached 4°C. The labelled bubble substrate was then stored at -20°C.

2.3 Alkaline cleavage deamination assay

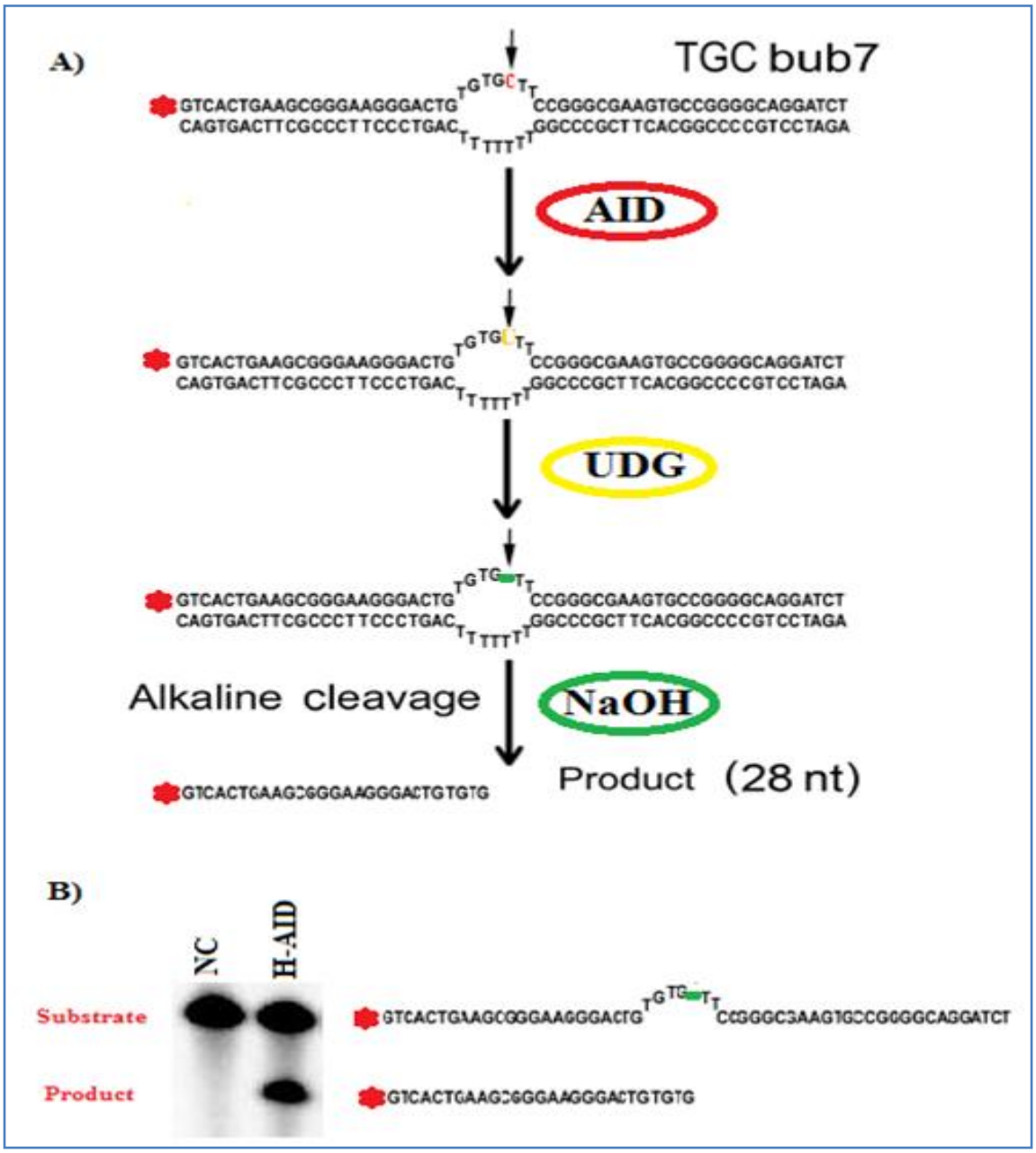
This assay has already been described (Sohail et al., 2003; Larijani et al., 2007; Dancyger et al., 2012) and was used to measure and quantify the catalytic activity of AID.

A schematic representation of the alkaline cleavage deamination assay is shown in figure 6. The deamination activity of AID was tested over a wide range of temperatures in order to precisely determine the optimal temperatures and thermosensitivity patterns of each wild type AID along with the hybrid AID enzymes. Briefly, 0.15-300 fmol/μL of the labelled bubble substrate was incubated with 0.1-0.25 μg GST-AID enzymes (to deaminate the cytidine within the hot spot motif WRC) in 100 mM phosphate buffer pH 7.3 in a total volume of 10μL per reaction tube, over a broad range and fine increments of temperatures. Negative controls were reactions with substrate and no enzymes, and positive controls were reactions containing enzymes previously purified and known to be active. The incubation time ranged from 2 to 6 hours followed by immediate increase in the temperature to 85°C for 30 minutes (then returned to 37°C for 10 min) to irreversibly deactivate AID so as to ensure that each AID enzyme was incubated only at the intended temperature before moving on to the next step which takes place at 37°C. A master mix of 7.8 μL H2O, 2 μL 10X UDG buffer and 0.2 μL UDG enzymes (NEB, Ipswich, MA, USA) was incubated with the previous reaction for 30 minutes to excise the uridine. The total reaction volume is now 20 ul. 2 μl of 2 M NaOH (fresh; 200 mM final) was then added and temperature was increased to 96°C for 10 minutes to cleave the alkali-labile abasic site. Immediately before loading the gel, 8 μL of formaldehyde loading dye were added to each reaction and incubated at 96°C for 5 minutes. Samples were loaded on a 14% polyacrylamide denaturing gel (1X TBE, 25% Formamide, 14% acrylamide:bisacrylamide [19:1], 7M Urea) and run at 200-300 V until dye front travelled approximately halfway down the gel. The resolved gels were exposed to a blanked Kodak

Storage Phosphor Screen GP (Bio-Rad) and imaged using a PhosphorImager (Bio-Rad, Hercules, CA, USA). Quantitation was done using an Image Lab™ 4.1 and Quantity One 1-D Analysis Software (Bio-Rad, Hercules, CA).

2.4 Data collection, quantification and analysis

Image Lab™ 4.1 analysis Software (Bio-Rad) was used to measure and compare the densitometry of the deaminated product band and the substrate band. Using these values, the relative amount of deaminated product (as a percentage of the total non-deaminated and deaminated bands) was derived. Background values are automatically determined by Image Lab, and were also manually determined from negative control lanes and subtracted from reaction lanes. For independent confirmation, each individual lane of an alkaline cleavage was also quantified three times using Quantity One 1-D Analysis software (Bio-Rad, Hercules, CA) to ensure the accuracy of data analysis. Experiments were repeated at least 2 and up to 6 times to ensure reproducibility of results. GraphPad Prism software (GraphPad, San Diego, CA, USA) was used to graph and analyze the obtained data.



(Figure A was modified from Abdouni et al., 2013)

Figure 6: Schematic representation of alkaline cleavage deamination assay.

Figure 6: Schematic representation of alkaline cleavage deamination assay.

A) Typical TGC bubble substrate used in the *in vitro* alkaline cleavage deamination assay. The red C within the hot spot motif TGC indicates the target cytidine. Asterisk (*) indicates the radioactively labeled strand. During the alkaline cleavage assay, AID deaminates the cytidine within the TGC motif turning it into uridine. The uridine is then excised by the addition of UDG enzyme producing an abasic site that is alkali-labile. This alkali-labile abasic site can be cleaved by adding NaOH to generate a cleaved product. By measuring and comparing the quantity of the cleaved product versus the non-cleaved substrate, we can precisely measure the deamination activity and then identify the thermosensitivity pattern of each AID enzyme. B) A result of alkaline cleavage assay showing the deamination activity of Human AID on TGCbub7 substrate versus the negative control sample (substrate without adding AID).

3. Results

3.1 Human AID has a distinct thermosensitivity and adaptive temperature from bony fish AIDs, but neither the catalytic nor N-terminal region of AID plays a significant role in these adaptive changes

Our lab recently showed that AID from various species have disparate catalytic efficiencies (deamination rates), and that the catalytic efficiency of AID from each species is "optimal" at natural physiological temperatures and is diminished at other temperatures (Dancyger et al., 2012). Specifically, when tested at three different temperatures, human AID displayed "optimal" deamination activity at 37°C compared to 25°C and 18°C, respectively. In contrast, bony fish AIDs (e.g. zebrafish AID, and catfish AID) exhibited reduced deamination rates at 37°C, but increased activities at 25°C and 18°C. However, the molecular basis and the underlying causes of the thermosensitivity of AID among various species are not known. Since the structure of AID has not yet been elucidated, we hypothesized that understanding this biochemical property can potentially reveal novel structure-function aspects. We proceeded to test our specific hypothesis that the thermosensitivity of AID is a characteristic encoded within a specific region of the enzyme, which can be transferred from one species' AID to another by swapping said region. Alternatively, this property might be attributed to the overall 3-dimensional structure of AID in a way that requires residues located throughout the enzyme rather than a region that is continuous in the primary structure. To test this hypothesis, the primary structure of AID was divided into three linear regions based on the position of the catalytic region (the region that contains the Zn-coordinating residues) and the

basic similarity of AID to other APOBEC family members. We defined the three linear regions of AID as the N-terminal, the catalytic (Zn-coordinating), and the C-terminal (figure 7). Hybrid AID enzymes involving portion(s) from human AID swapped with other AID portion(s) from either zebrafish or channel catfish or takifugu AID were constructed in order to narrow down and identify which region of AID is specifically responsible for this biochemical property. The three letter description of each hybrid denotes the N-, catalytic, and carboxyl-terminal regions, respectively. Our hybrid AID enzymes were designed as follows: HDH and HIH AIDs are human AIDs (Hs-AID) with the catalytic regions of zebrafish AID (Dr-AID) and catfish AID (Ip-AID), respectively; HII AID is an Ip-AID with the N-terminal region of Hs-AID. These hybrid AIDs along with the aforementioned wild-type AID enzymes were independently cloned into our GST-fusion pGEX-5-3 expression vectors, expressed in *E.coli* and purified on Glutathione agarose beads to produce GST-tagged AID as previously described (Larijani et al., 2007). SDS PAGE was utilized to determine the purity of each AID enzyme used in the following experiments. The deamination rate and the thermosensitivity of each purified AID were tested using the alkaline cleavage assay. In this approach, each purified AID was incubated with radioactively labeled DNA substrate that has the AID-favored hot spot motif TGC within partially single-stranded bubble sequences (figure 9). This bubble substrate with 7 nt long was shown to be an optimal target for AID deamination (Larijani et al., 2007). The incubation was carried out at three different temperatures 37°C, which is the normal human body temperature; 25°C and 18°C, which represent a range of temperatures where some bony fish (e.g. Zebrafish and channel catfish) naturally live (Shrable et al., 1969; Chen 1976; Spence et al., 2006). During

incubation, AID deaminates the cytidine within the TGC motif turning it into uridine. The uridine is then excised by the addition of UDG enzyme producing an abasic site that is alkali-labile (Larijani & Martin, 2007). NaOH is then added to cleave the alkali-labile abasic site generating a cleaved product (figure 6). By measuring and comparing the quantity of the cleaved product versus the non-cleaved substrate, we can precisely measure the deamination activity and then identify the thermosensitivity pattern of each AID enzyme. As expected, Dr-AID and Ip-AID demonstrated higher deamination rates at 25°C than at 37°C (in average of 6 experiments: Dr-AID 59% at 18°C, 69% at 25°C, and 26% at 37°C; Ip-AID 11% at 18°C, 17% at 25°C, and 4% at 37°C), whereas Hs-AID was more catalytically active at 37°C than at 25°C and 18°C (in average of 6 experiments: 39% at 37°C *vs.* 21%, 14% at 25°C and 18°C, respectively). Both Dr-AID and Ip-AID (regardless of the difference in their activity levels) exhibit the same general trend of thermosensitivity in that they have diminished deamination rate at 37°C and maximum activity at 25°C (figure 10). Quite surprisingly, both HDH and HIH hybrid AIDs have enhanced catalytic rates at 37°C rather than at lower temperatures (In average of 6 experiments: HDH 17% at 37°C *vs.* 9%, 7% at 25°C and 18°C, respectively; HIH 25% at 37°C *vs.* 21%, 19% at 25°C and 18°C, respectively), and each of them displayed the same thermosensitive modality as Hs-AID (figure 11-A). In contrast, the HII hybrid AID enzyme revealed similar thermosensitivity pattern to Ip-AID (figure 11-B). We noticed that even though the catalytic region of Hs-AID was swapped once with that of Dr-AID (e.g. HDH) and with Ip-AID (e.g. HIH), this exchange does not seem to have an effect on the thermosensitivity pattern of Hs-AID. In addition, interchanging the N-terminal end of the catfish AID with its equivalent of human AID (e.g. HII) does not transfer the

thermosensitivity from the latter to the former. Therefore, we concluded that neither the catalytic nor the N-terminal region of AID seems to play a noteworthy role in influencing the thermosensitivity of AID from human and bony fish.

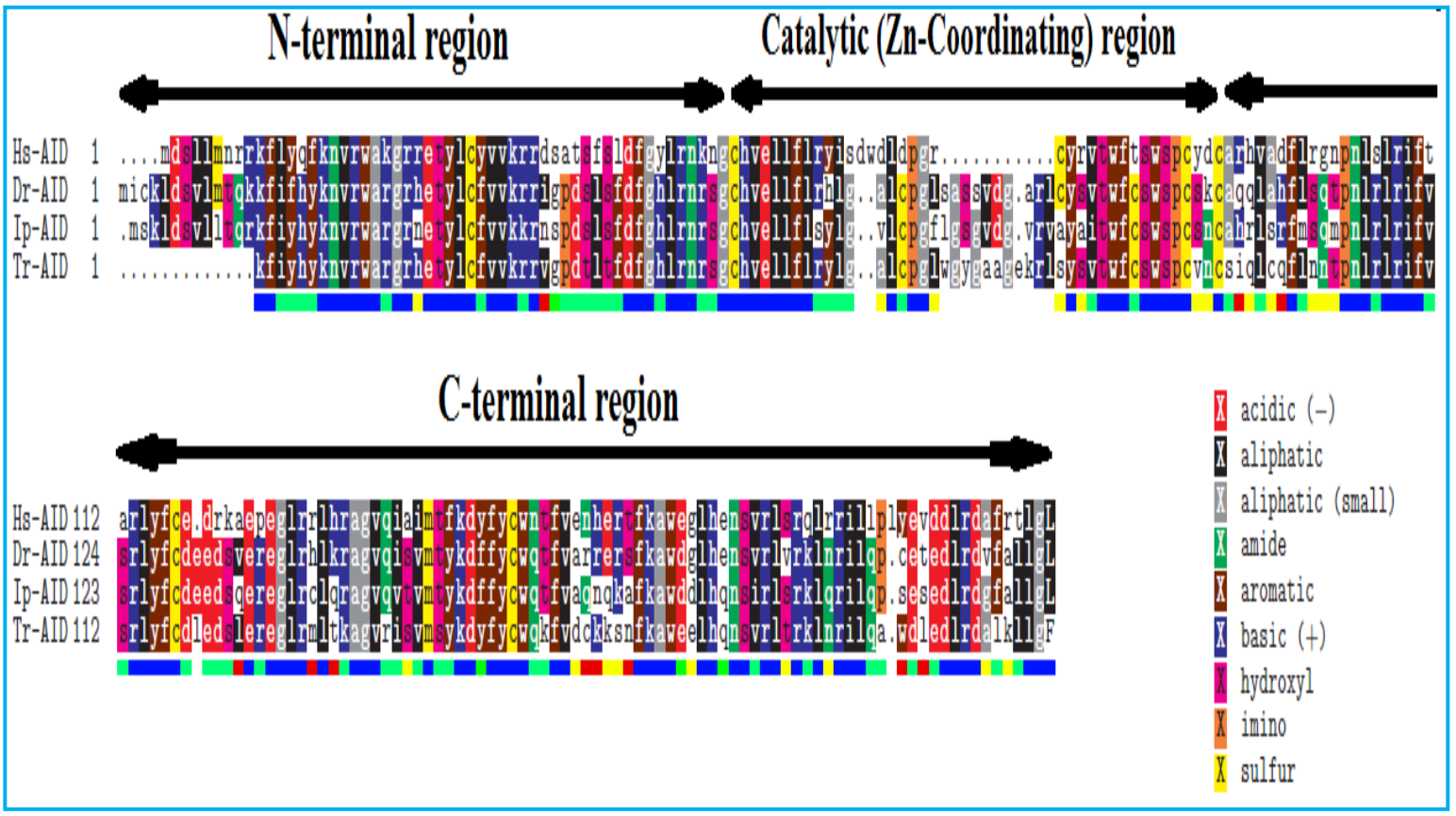


Figure 7: Amino acid sequence alignment of human versus bony fish AID.

Figure 7: Amino acid sequence alignment of human versus bony fish AID.

Amino acid sequences of human AID were compared with their equivalents sequences of zebrafish, catfish, and takifugu AIDs. Hs-AID, Dr-AID, Ip-AID, and Tr-AID stand for human, zebrafish, catfish, and pufferfish (takifugu) AIDs, respectively. Residues are colored according to the side chain. Blue underlining indicates identical residues, green indicates similar residues, and red indicates disparate residues. The boundaries of the carboxyl, catalytic, and N-terminal regions are indicated by the arrows. These amino acid sequences are obtained from the published sequences of AID (*Homo sapien*: AA95406.1 for human AID, *Danio rerio*: NP001008403 for zebrafish AID, *Ictalurus punctatus*: AAR7544 for catfish AID, *Takifugu rebripes*: AAU95747 for pufferfish AID), and aligned using Strap Alignment software (http://www.bioinformatics.org/strap/).

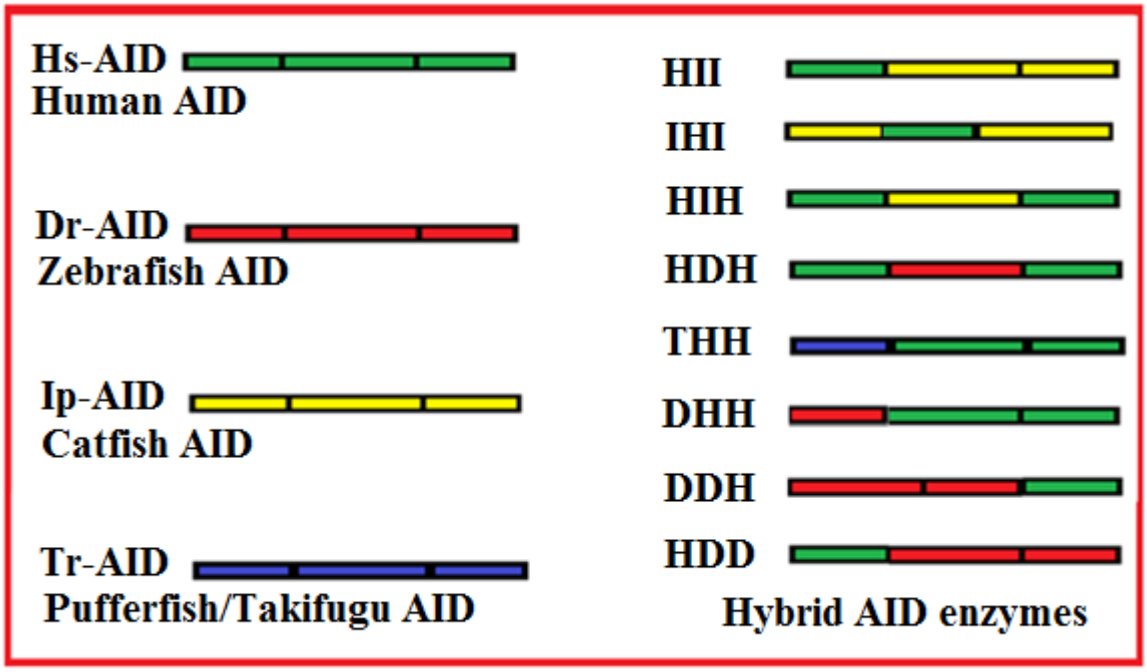


Figure 8: Schematic representation of hybrid AID enzymes.

Figure 8: Schematic representation of hybrid AID enzymes.

Hybrid AID enzymes involving different region(s) from human AID linked with other AID region(s) from either zebrafish or catfish or takifugu AIDs. These hybrid AID enzymes were designed and constructed in order to determine which hybrid AID with an altered thermosensitivity and specify which region would transfer the thermosensitivity from one species to another. The three letter description of each hybrid AID indicates the N-, catalytic and carboxyl-terminus regions, respectively. The green, red, yellow, and blue bars denote Human, zebrafish, catfish, and takifugu AIDs, respectively.

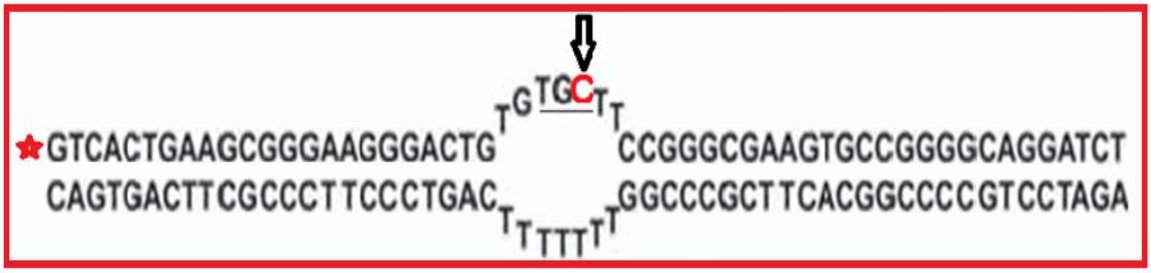


Figure 9: Partially single-stranded bubble substrate used in alkaline cleavage deamination assay.

Figure 9: Partially single-stranded bubble substrate used in alkaline cleavage deamination assay.

Radioactively labeled DNA substrate with 56 nt in length containing a 7 nt long single-stranded bubble region that has hot spot motif TGC for AID deamination. During the alkaline cleavage assay, AID deaminates the cytidine nucleotide located within TGC motif to uridine. The targeted cytidine is colored in red and the radioactively labeled strand is denoted by (*).

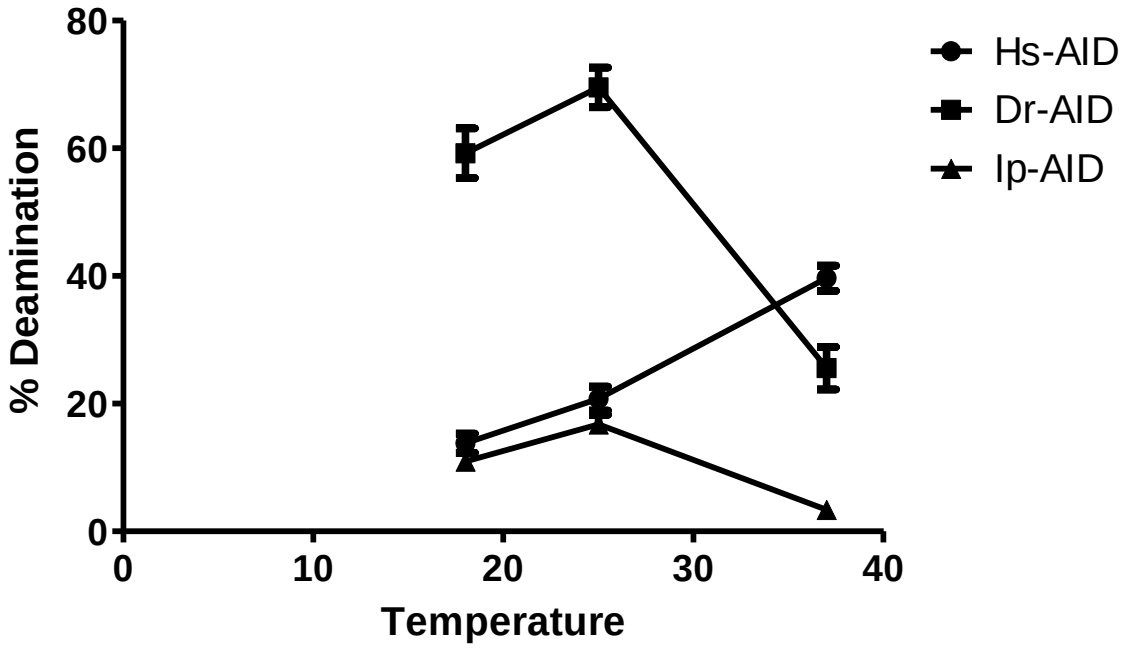


Figure 10: Comparison of thermosensitivity patterns of Hs-AID Dr-AID, and Ip-AID at three different temperatures.

Figure 10: Comparison of thermosensitivity patterns of Hs-AID Dr-AID, and Ip-AID at three different temperatures.

The thermosensitivity pattern of human AID was compared with zebrafish and catfish AIDs at 37°C, 25°C, and 18°C. Each AID enzyme was incubated for 2 hours with 50 fmol of TGCbub7 substrate. The quantity of the deamination by each AID at each temperature was measured using Image Lab™ 4.1 software and confirmed by Quantity One 1-D analysis software. The percentage of the deaminated products was plotted. Dr-AID and Ip-AID exhibited different levels of deamination; nevertheless they have comparable thermosensitivity patterns. Both of Dr-AID and Ip-AID (regardless of the difference in their activity levels) have maximum activity at 25°C, whereas the enzymatic activity of Hs-AID declined as temperature was reduced.

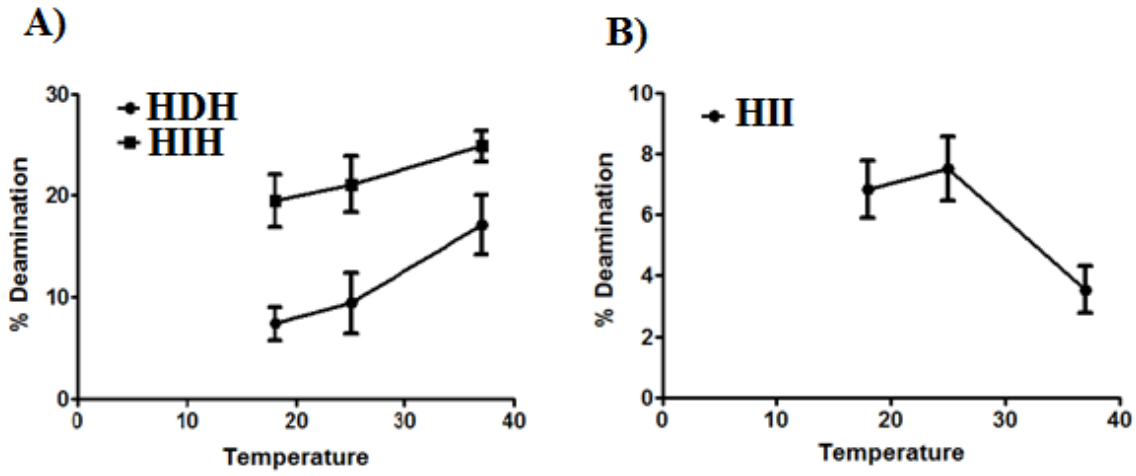


Figure 11: Comparison of thermosensitivity patterns of HDH, HIH, and HII at three different temperatures.

Figure 11: Comparison of thermosensitivity patterns of HDH, HIH, and HII at three different temperatures.

The thermosensitivity patterns of the hybrid AID enzymes: HDH, HIH (figure 10-A), and HII (figure 10-B) at 37°C, 25°C, and 18°C were compared. Both HDH and HIH are Hs-AIDs with catalytic region of Dr-AID and Ip-AID, respectively. HII is a Ip-AID with N-terminal region of Hs-AID. Each of these hybrid AID enzymes was incubated for 2 hours with 50 fmol of TGCbub7 substrate. Hs-AID, Dr-AID, and Ip-AID were used as positive controls in this experiment, but were represented in a separate graph. The amount of the deamination by each hybrid AID at each temperature was measured using Image Lab™ 4.1 software and confirmed by Quantity One 1-D analysis software. The percentage of the deaminated products was plotted. HDH and HIH AIDs displayed comparable thermosensitivity patterns to that of Hs-AID, whereas the thermosensitivity pattern of HII AID was equivalent to that Ip-AID.

3.2 AID activity is not always optimal at native physiological temperatures

Even though testing the wild type AIDs (Hs-AID, Dr-AID, and Ip-AID) at three different temperatures substantiated the idea that AID from various species have disparate thermosensitivity patterns, the *exact* optimal temperature for each AID enzyme was still unknown. For that reason, the wild type AIDs were examined over a wider range and smaller increments of temperatures to more precisely determine the optimal temperature conditions for each AID enzyme. Initially, Hs-AID, Dr-AID, and Ip-AID were each tested at 12°C, 14°C, 16°C, 18°C, 20°C, 22°C, 25°C, 27°C, 30°C, 32°C, 34°C, 37°C, 39°C, and 41°C. Intriguingly, the optimal temperature for Hs-AID *in vitro* was not 37°C as formerly assumed; instead, in average of 5 independent experiments, Hs-AID was 1.6 fold more catalytically active at 32°C than at 37°C, whereas the precise optimal temperatures for Dr-AID and Ip-AID was 27°C and 22°C, respectively (figure 12). This finding enabled us to more precisely define whether a hybrid AID is more similar to human or fish in terms of its thermosensitivity. Specifically, we set the definition that *regardless of the absolute activity level compared to other enzymes,* if a hybrid AID was optimal at 25°C (+/- 2°C) then it will be categorized "fish-like" in its thermosensitivity. If on the other hand the hybrid AID was optimal at 32°C (+/- 2°C), then it will be categorized as "human-like" or warm-adapted.

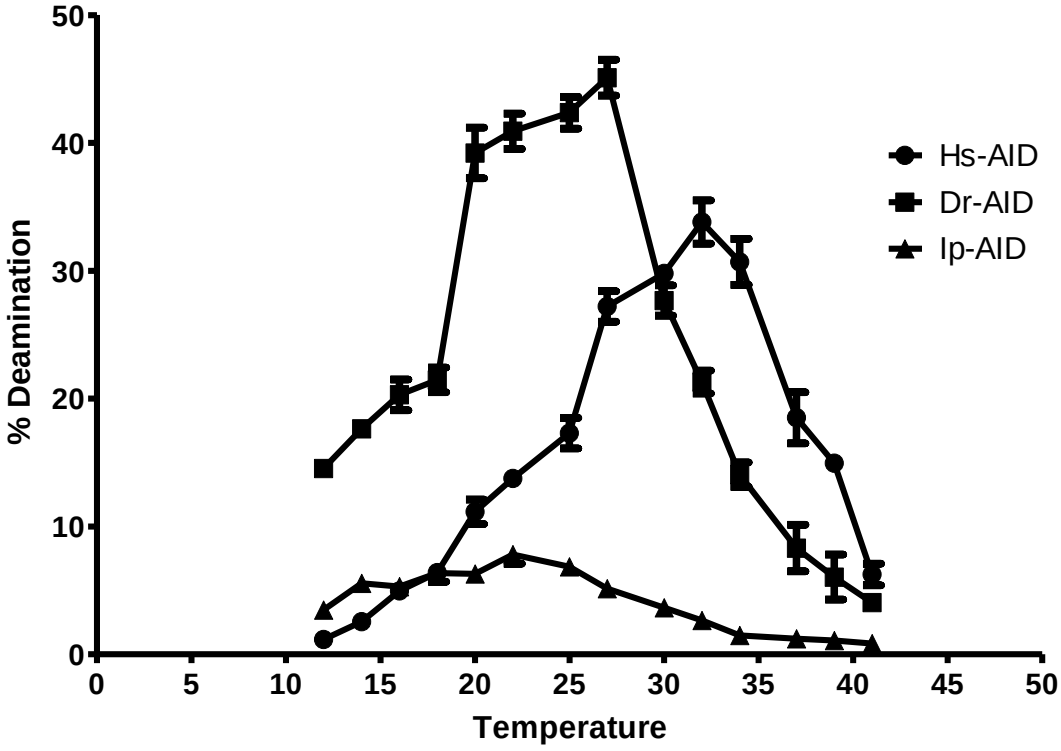


Figure 12: Comparison of optimal temperatures and thermosensitivity patterns of Hs-AID, Dr-AID, and Ip-AID over a wide range of temperatures.

Figure 12: Comparison of optimal temperatures and thermosensitivity patterns of Hs-AID, Dr-AID, and Ip-AID over a wide range of temperatures.

The deamination activities and thermosensitivity patterns of Hs-AID, Dr-AID, and Ip-AID were compared over a wide range and small increments of temperature from 12°C to 41°C. The optimal temperatures for each of these AID enzymes were determined and plotted. 50 fmol of TGCbub7 substrate was incubated with each AID for 2 hours at 12°C, 14°C, 16°C, 18°C, 20°C, 22°C, 25°C, 27°C, 30°C, 32°C, 34°C, 37°C, 39°C, and 41°C. The reaction temperature was immediately increased to 85°C for 30 minutes (then decreased to 37°C for 10 minutes) to irreversibly deactivate AID and so as to ensure that each AID enzyme was incubated only at our intended temperature before moving on to the next step which takes place at 37°C. The percentage of the each deaminated product was plotted. The optimal temperature for Hs-AID is 32°C, whereas Dr-AID and Ip-AID are optimal at 27°C and 22°C, respectively.

3.3 Optimal temperatures of hybrid AID enzymes.

Following our approach of precisely determining the optimal temperatures of each AID enzyme using finer increments, more hybrid AID enzymes were designed (figure 8), and tested over a wide range of temperatures in order to identify the optimal temperature for each hybrid AID. In addition, we endeavoured to determine which region specifically is responsible for the thermosensitivity of AID. Different regions from human AID were pieced together with other region(s) from either the zebrafish or takifugu AIDs as following: DHH, THH, HDD and DDH AID. We utilized the N-terminal region from Tr-AID (as another bony fish AID) and swapped it with the catalytic and the C-terminal region of Hs-AID to generate THH AID. This newly constructed hybrid AID was used to further confirm our conclusion up to this point, that the N-terminal region has no major role in influencing the thermosensitivity of AID. Because this was our first analysis of Tr-AID and we did not know its sequence preference, we chose to test it on 3 different WRC motif-containing substrates: TGCbub7, AGCbub7 and TACbub7. Despite the fact that Tr-AID exhibited very low deamination activity *in vitro*, we found that it has the same general thermosensitivity pattern as Dr-AID and Ip-AID and also more preference for the TGCbub7 substrate than AGCbub7 and TACbub7 substrates, respectively (figure 13)**.** In agreement with our previous findings, both HDH and HIH hybrid AIDs demonstrated comparable optimal temperatures and thermosensitivity patterns to that of Hs-AID when they were tested at various temperatures ranging from 12°C to 41°C (30°C and 32°C were the optimal temperatures for HIH, HDH, respectively, *vs*. 32°C for Hs-AID) (figure 14). This result confirmed that the catalytic region of AID has no significant effect on determining the thermosensitivity pattern. On the other hand, HII AID displayed

low catalytic efficiency, but nevertheless exhibited a thermosensitivity pattern identical to that of Ip-AID (22°C is the optimal temperature for HII *vs.* 22°C for Ip-AID) (figure 15). In addition, four independent preparations of THH and DHH (THH2, THH8, DHH2, and DHH7) were tested independently at different temperatures ranging from 12°C to 41°C. None of THH and DHH hybrid AIDs showed a disparate thermosensitive pattern from Hs-AID (figure 16), although the N-terminus of each of these hybrid AIDs does not belong to human AID (the optimal temperature was 30°C for DHH2, 31°C for DHH7, 33°C for THH2, and 32°C for THH8 *vs.* 32°C for Hs-AID). These observations further strengthened our conclusion thus far that the N-terminus is inconsequential for thermosensitivity. Moreover, in confirmation of our previous results, HDD hybrid AIDs (HDD8 & HDD9) exhibited optimal temperatures and thermosensitivity patterns corresponding to Dr-AID; notwithstanding that they have the N-terminus of human AID (28°C and 26°C for HDD8 and HDD9 *vs.* 27°C for Dr-AID) (figure 17). In summary, all our data up to now endorse the notion that the N-terminal and the catalytic region of AID do not transfer the thermosensitivity pattern of AID from human to bony fish or vice versa.

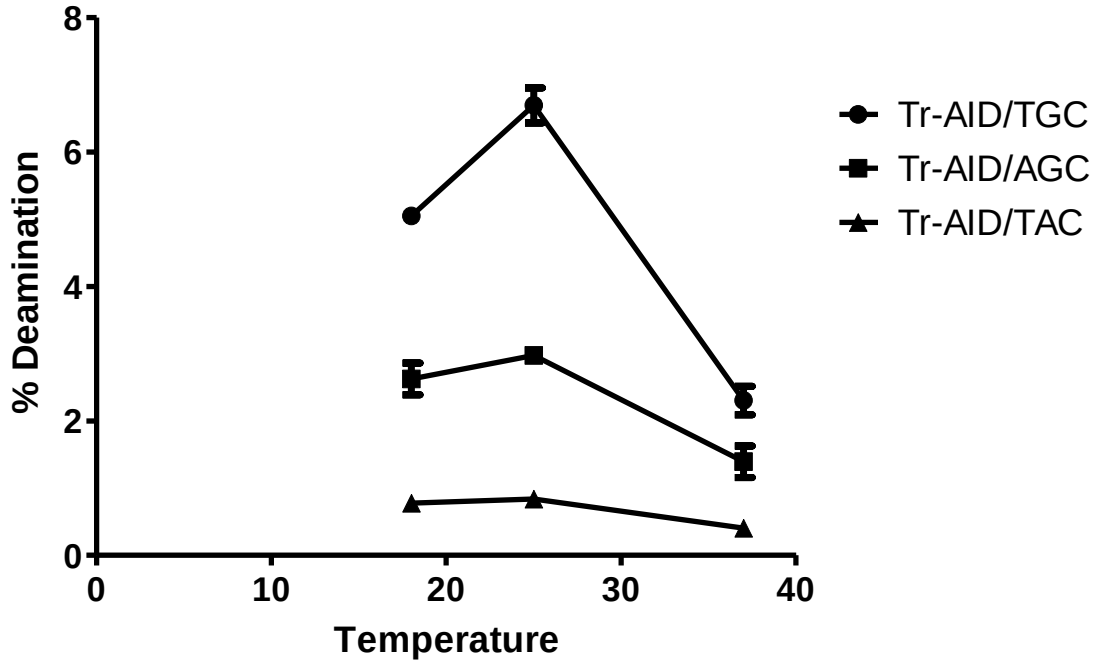


Figure 13: Temperature and substrate preference of Tr-AID.

Figure 13: Temperature and substrate preference of Tr-AID.

The deamination activity of takifugu AID (Tr-AID) was tested at three different temperatures (37°C, 25°C, and 18°C) on three different substrates (TGCbub7, AGCbub7, and TACbub7). 50 fmol of each substrate was incubated with Tr-AID for 6 hours. Since this purified Tr-AID has low enzymatic activity *in vitro*, the incubation time was increased from 2 to 6 hours. The amounts of the deaminated products produced by Tr-AID at each temperature with each substrate were plotted. TGCbub7 is the optimal target substrate for deamination *in vitro* by Tr-AID. Image Lab™ 4.1 software and Quantity One 1-D analysis software programs were used to measure the quantity of the deamination activity generated by Tr-AID.

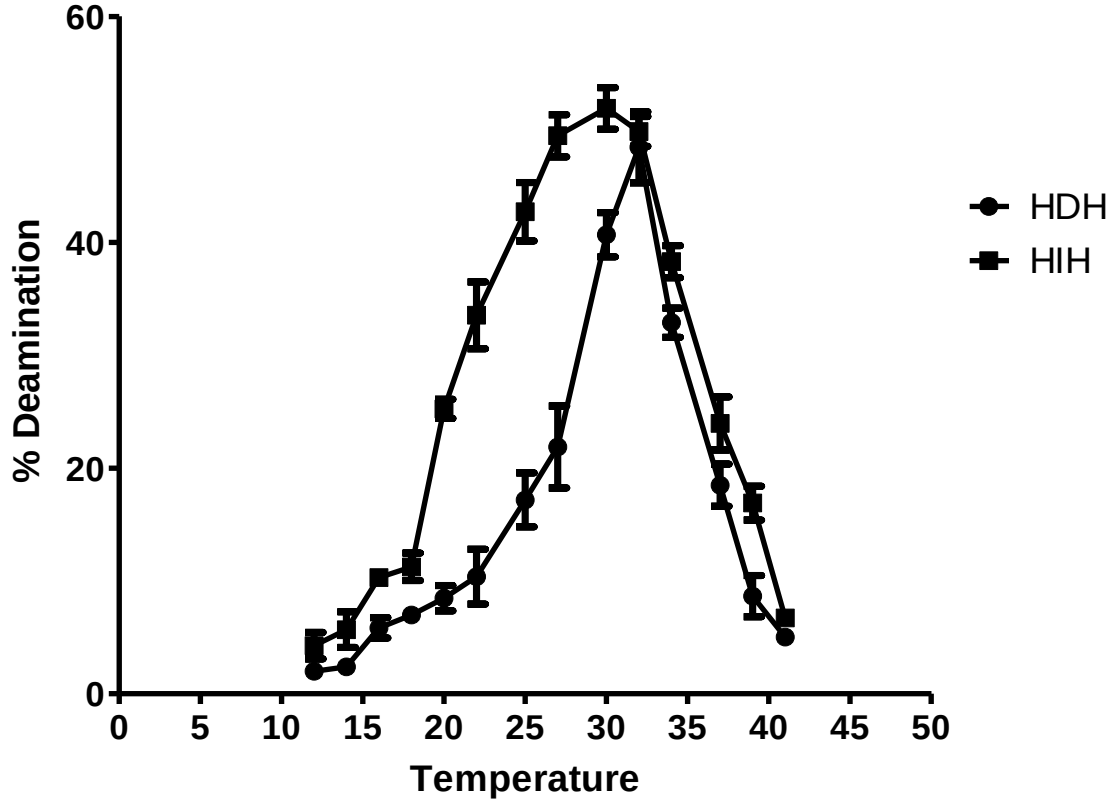


Figure 14: Comparison of thermosensitivity patterns and optimal temperatures of HDH and HIH hybrid AIDs.

Figure 14: Comparison of thermosensitivity patterns and optimal temperatures of HDH and HIH hybrid AIDs.

The optimal temperatures and thermosensitivity patterns of HDH and HIH AIDs were compared over a wide range and fine increments of temperature from 12°C to 41°C. 50 fmol of TGCbub7 substrate was incubated with each of the hybrid AIDs for 2 hours at 12°C, 14°C, 16°C, 18°C, 20°C, 22°C, 25°C, 27°C, 30°C, 32°C, 34°C, 37°C, 39°C, and 41°C. The optimal temperatures for HDH and HIH AIDs were determined (32°C for of HDH *vs.* 30°C for HIH). The percentage of the deaminated products plotted.

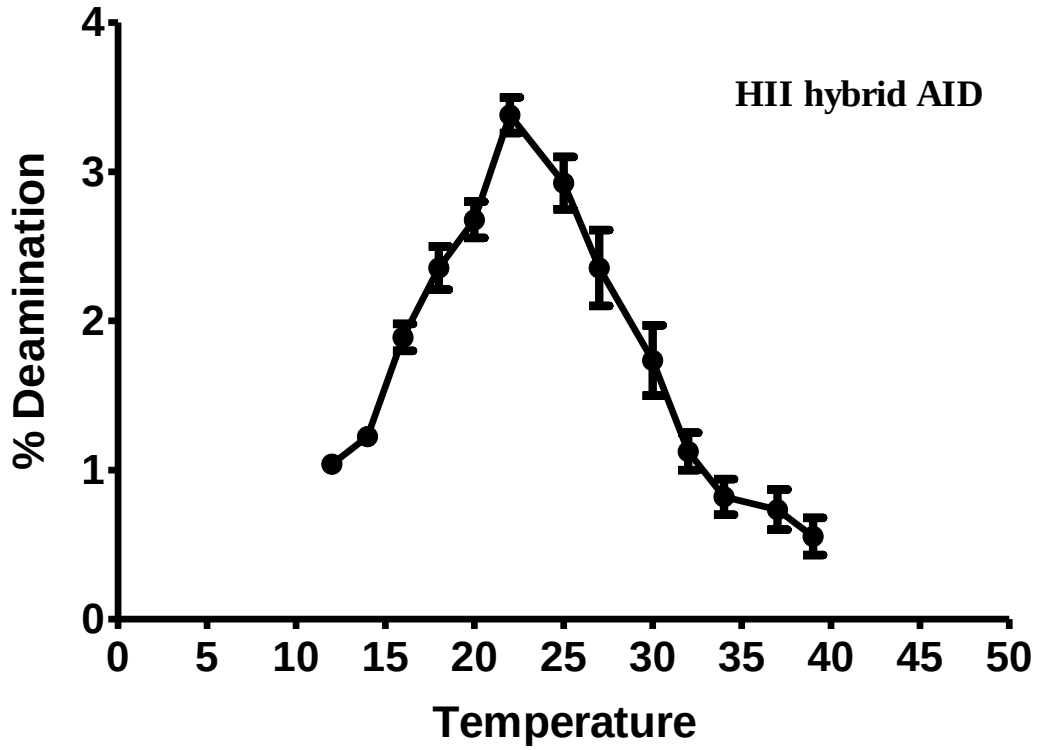


Figure 15: The thermosensitivity pattern and optimal temperature of HII hybrid AID.

Figure 15: The thermosensitivity pattern and optimal temperature of HII hybrid AID.

The deamination activity of HII hybrid AID was measured at temperatures ranging from 12°C to 41°C. 50 fmol of TGC substrate was incubated for 2 hours with HII AID at 12°C, 14°C, 16°C, 18°C, 20°C, 22°C, 25°C, 27°C, 30°C, 32°C, 34°C, 37°C, 39°C, and 41°C. The percentage of the deaminated products by HII at each temperature was plotted. The overall thermosensitivity pattern as well as the optimal temperature of HII was determined and illustrated. The optimal temperature of HII is identical with that of catfish AID (22°C is the optimal temperature for HII *vs.* 22°C for Ip-AID).

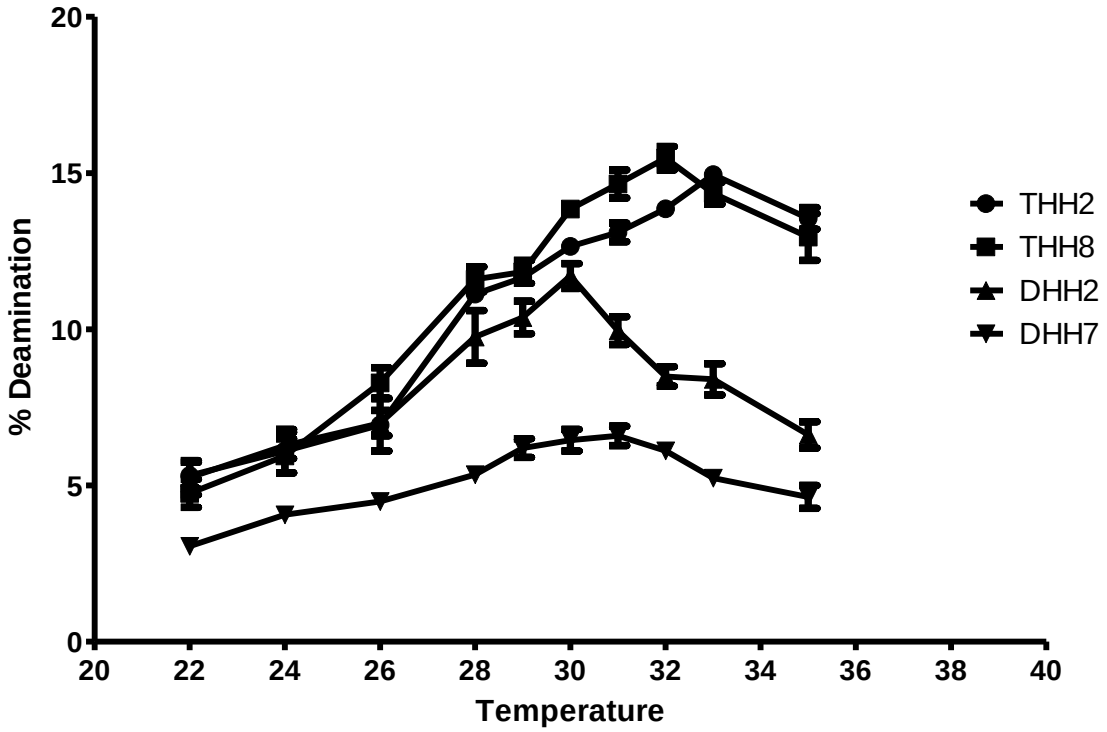


Figure 16: Comparison of optimal temperatures and thermosensitivity patterns of four independent preparations of THH and DHH hybrid AIDs.

Figure 16: Comparison of optimal temperatures and thermosensitivity patterns of four independent preparations of THH and DHH hybrid AIDs.

Each of the four independent preparations of THH and DHH (THH2, THH8, DHH2, and DHH7) was tested on TGCbub7 substrate for 2 hours at 12°C, 14°C, 16°C, 18°C, 20°C, 22°C, 25°C, 27°C, 30°C, 32°C, 34°C, 37°C, 39°C, and 41°C. Regardless of the disparate levels of the deamination activity generated by each of these hybrid AIDs, the optimal temperatures for all of them were consistent with that of Hs-AID (30°C for DHH2, 31°C for DHH7, 33°C for THH2, and 32°C for THH8 *vs.* 32°C for Hs-AID). The percentage of the deaminated products was graphed and the overall thermosensitivity pattern for each independent AID preparation was presented.

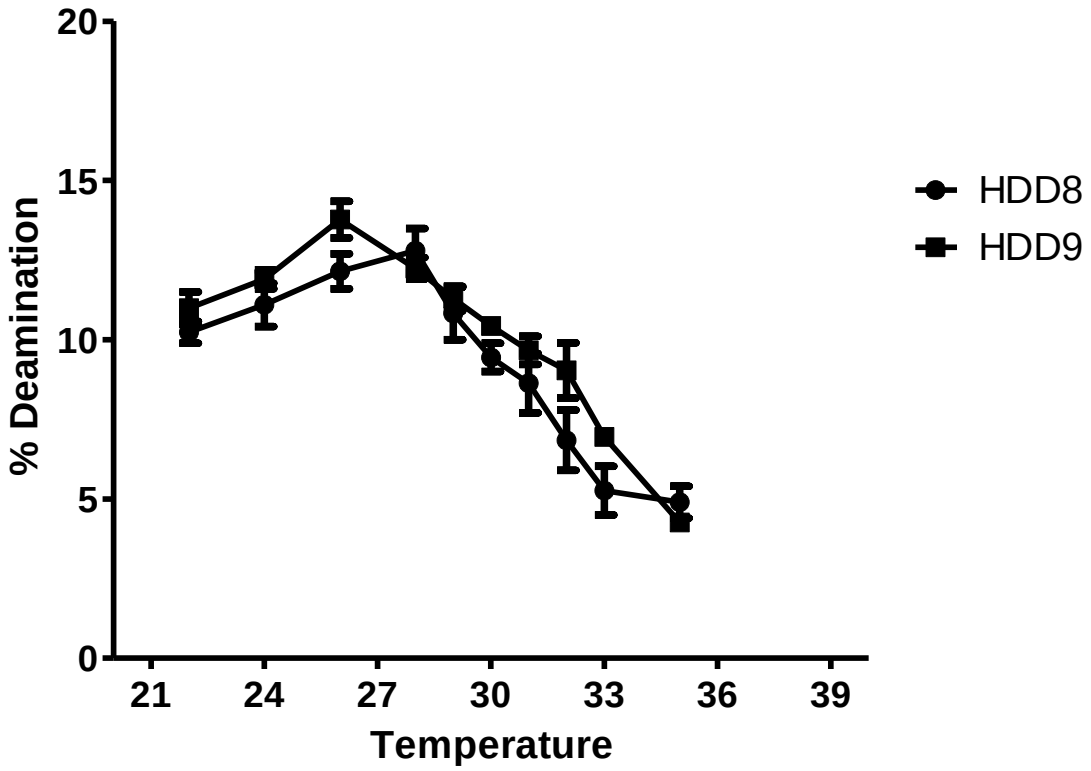


Figure 17: Comparison of optimal temperatures and thermosensitivity patterns of two independent preparations of HDD hybrid AID (HDD8 & HDD9).

Figure 17: Comparison of optimal temperatures and thermosensitivity patterns of two independent preparations of HDD hybrid AID (HDD8 & HDD9).

The thermosenitivity patterns of two independent preparations of HDD (HDD8, HDD9) were compared. Each of the HDD preparations was incubated with 50 fmol of TGCbub7 substrate for 2 hours at 12°C, 14°C, 16°C, 18°C, 20°C, 22°C, 25°C, 27°C, 30°C, 32°C, 34°C, 37°C, 39°C, and 41°C.The percentage of the deaminated products by HDD8 and HDD9 AIDs at each temperature was plotted and the thermosensitivity patterns for both of them were graphed. The optimal temperatures of HDD8 and HDD9 are similar to that of Dr-AID (28°C and 26°C for HDD8 and HDD9 *vs.* 27°C for Dr-AID).

3.4 The C-terminal region plays an important role in the thermosensitivity of AID

After excluding the possibility of significant roles of the N-terminal and catalytic regions in influencing AID thermosensitivity, our attention was drawn to the C-terminal region of AID. At that point, four independent preparations of DDH hybrid AID (DDH5, DDH6, DDH6', and DDH9) were purified (as already described) and tested along with the previously constructed hybrid AIDs over a wide range of temperatures ranges from 12°C to 41°C. Wild-type Hs-AID and Dr-AID were used as controls in this experiment. Although the four independent preparations of DDH AID displayed various levels of the deamination activity (which might be attributed to varying purification and expression conditions), they exhibited near identical thermosensitivity patterns closer to that of Hs-AID (31°C for DDH5, DDH6, 30°C for DDH6'and DDH9 *vs.* 32°C for Hs-AID) (figure 18). From this data, and previous work excluding the N- and catalytic regions, we conclude that the C-terminal region is the most significant part of AID in terms of containing the residue(s) that is/are responsible for thermosensitivity of AID. It is worth mentioning that the same conclusion can be drawn if particular attention was paid to the C-terminal end of any of our hybrid AIDs. In other words, we noticed that all the hybrid AID enzymes that have the C-terminal region of human AID, regardless of the origin of the N-terminal and catalytic regions (e.g. HDH, HIH, THH, DHH, and DDH), demonstrated thermosensitivity patterns and optimal temperatures similar to that of Hs-AID (32°C, 30°C, 32°C, 31°C and 31°C for HDH, HIH, THH, DHH, and DDH, respectively, *vs.* 32°C for Hs-AID) (figure 11-A, 14, 16, and 18). In contrast, the hybrid AIDs that have a C-terminal region which belongs to Dr-AID (e.g. HDD) or Ip-AID (e.g. HII) exhibited thermosensitivity patterns and optimal temperatures paralleling those of

Dr-AID and Ip-AID, respectively (28°C and 26°C for HDD8 and HDD9, respectively, *vs.* 27°C for Dr-AID; 22°C for HII *vs.* 22°C for Ip-AID) (figure 11-B, 15, and 17). To date, all our findings support the notion that the thermosensitivity of AID is mostly a C-terminal region-based characteristic.

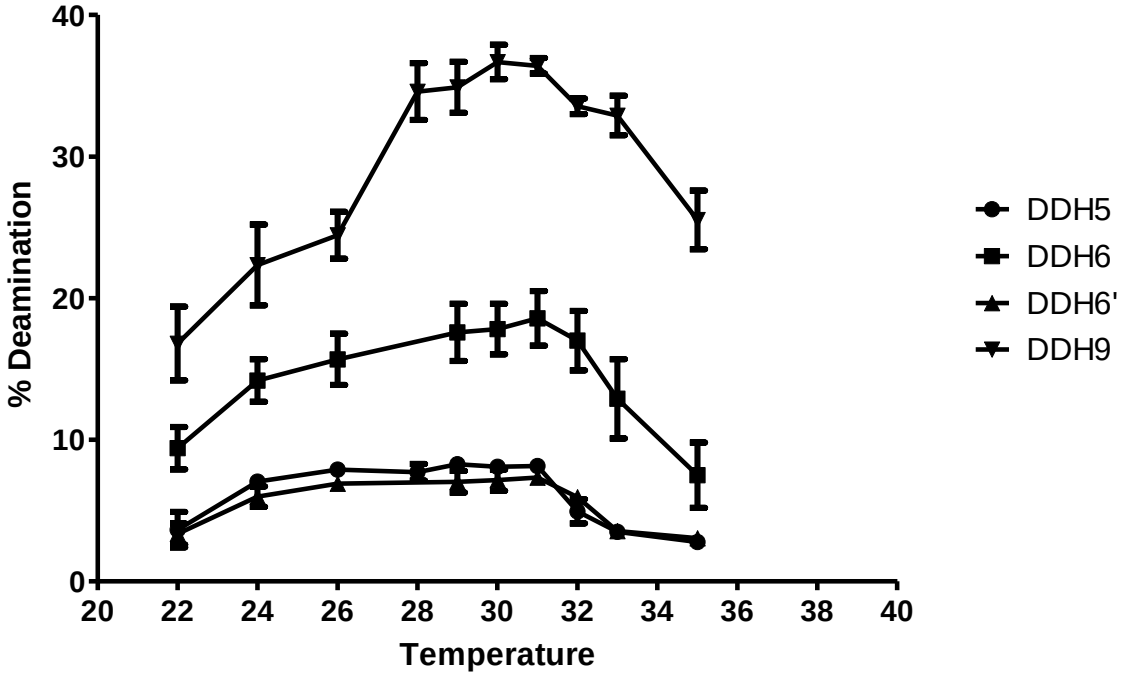


Figure 18: Comparison of optimal temperatures and thermosensitivity patterns of four independent preparations of DDH hybrid AID.

Figure 18: Comparison of optimal temperatures and thermosensitivity patterns of four independent preparations of DDH hybrid AID.

Each of the four independent preparations of DDH was incubated with 50 fmol of TGCbub7 substrate for 2 hours at 12°C, 14°C, 16°C, 18°C, 20°C, 22°C, 25°C, 27°C, 30°C, 32°C, 34°C, 37°C, 39°C, and 41°C. The percentage of the deaminated products at each temperature was plotted and the overall thermosensitivity patterns of DDH5, DDH6, DDH6' and DDH9 were graphed. Regardless of the different levels of the deamination activity of each preparation of DDH, the optimal temperatures for all of them were comparable with that of Hs-AID (31°C for DDH5, DDH6, 30°C for DDH6'and DDH9 *vs.* 32°C for Hs-AID). Image Lab™ 4.1 software and Quantity One 1-D analysis software programs were used to measure the quantity of the deamination activity generated by each AID enzyme.

3.5 Comparison of deamination kinetics of human AID at 32°C versus 37°C, and that of zebrafish AID at different temperatures

Since 37°C is the physiological human body temperature, most of the previously conducted studies on human AID were performed at this temperature. However, our recent findings provided convincing evidence that the optimal temperature *in vitro* for human AID to deaminate the cytidine in the hot spot motif TGCbub7 is 32°C instead of 37°C (the deamination activity of Hs-AID was 33% at 32°C *vs.* 19% at 37°C) (figure 12). In contrast, we showed that the optimal temperature for Dr-AID *in vitro* is 27°C. Previous experiments were carried out with a single concentration of enzyme and substrate and a single incubation time, which does not allow for evaluation of true enzyme velocity. Thus, to quantitatively evaluate the extent of enzyme velocity differences at optimal vs. sub-optimal temperatures, we measured the deamination kinetics of each of Hs-AID and Dr-AID at these temperatures. We reasoned that this would also serve to further confirm our previous findings. To this end, we tested the deamination kinetics and the enzymatic activity of Hs-AID on TGCbub7 substrate at 32°C, 37°C, and 41.5°C, whereas Dr-AID was tested at 16°C, 27°C, and 33°C. These different temperatures were particularly chosen to test Dr-AID, based on recorded observations of the natural water temperatures that ranged from 16.5°C to 33°C from nine different sites at which zebrafish were collected (Spence et al., 2006) and because our own work described above determined that the optimal temperature for zebrafish was 27°C.

The velocity of the reactions was measured as a function of substrate concentration over a 120 fold dilution range (0.25-30 nM). In agreement with our previous results, Hs-AID exhibited more robust enzymatic potential and greater velocity of product formation at 32°C than 37°C and 41°C (the greatest reaction velocities at 32°C, 37°C, and 41°C were 0.195, 0.099, and 0.024 fmol product/μgAID/min, respectively, at the point of maximal difference) (figure 19). On the other hand, the deamination kinetics of Dr-AID was tested on the same substrate (TGCbub7) at 16°C, 27°C, and 33°C. Consistent with what we found thus far, Dr-AID reached the highest velocity of product formation at 27°C (the greatest reaction velocities at 16°C, 27°C, and 31°C were 0.60, 1.30, and 0.66 fmol product/μgAID/min, respectively, at the point of maximal difference) (figure 20). In conclusion, measuring the deamination kinetics of Hs-AID and Dr-AID revealed that the optimal temperature for maximal enzyme velocity is 32°C and 27°C for Hs-AID and Dr-AID, respectively. Furthermore, as expected from previous work (Dancyger et al., 2012; Larijani & Martin, 2012), under their optimal temperature conditions, Dr-AID demonstrated more potent deamination efficiency than Hs-AID (45 *vs.* 33%) (figure 12).

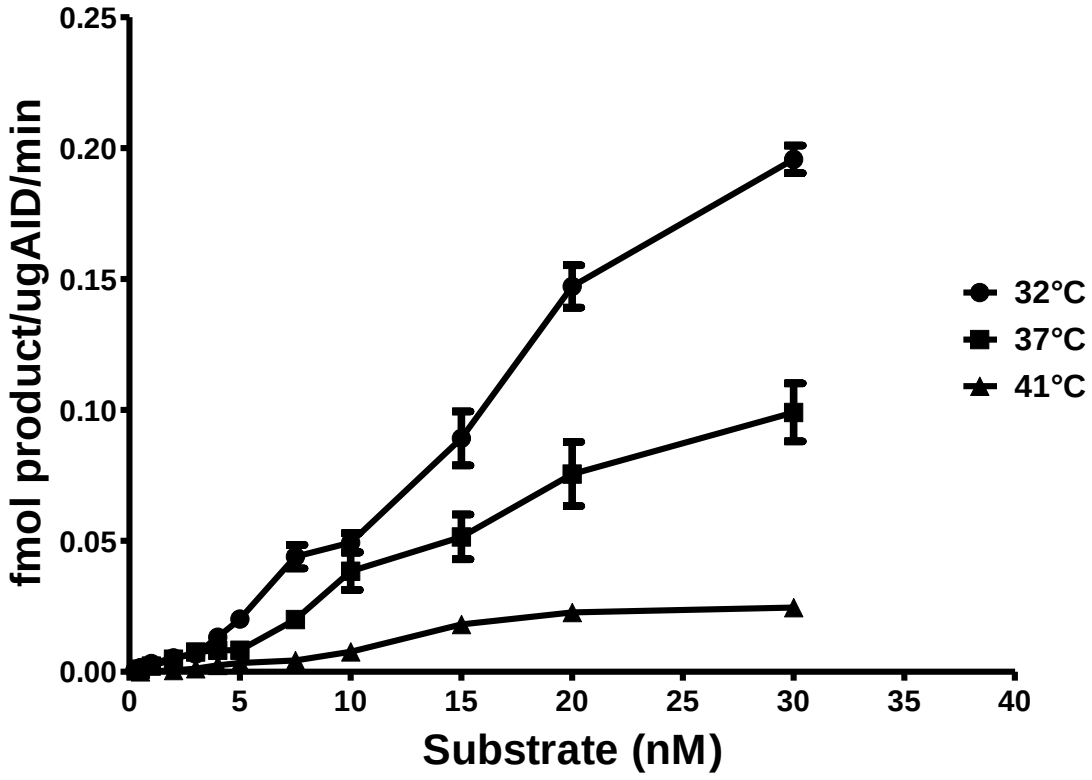


Figure 19: Comparison of deamination kinetics of Hs-AID at three different temperatures.

Figure 19: Comparison of deamination kinetics of Hs-AID at three different temperatures.

The deamination kinetics of Hs-AID on TGCbub7 substrate was compared at 32°C, 37°C, and 41°C. The velocity of the reactions was measured as a function of substrate concentration over a 120 fold dilution range (0.25-30 nM). The velocity of AID reaction is defined as the amount of deaminated product produced by a given amount of AID enzyme in a unit of time and plotted against substrate concentration. At 32°C Hs-AID reached its maximum enzymatic potential (the greatest reaction velocities at 32°C, 37°C, and 41°C were 0.195, 0.099, and 0.024 fmol product/ugAID/min, respectively, at the point of maximal difference).

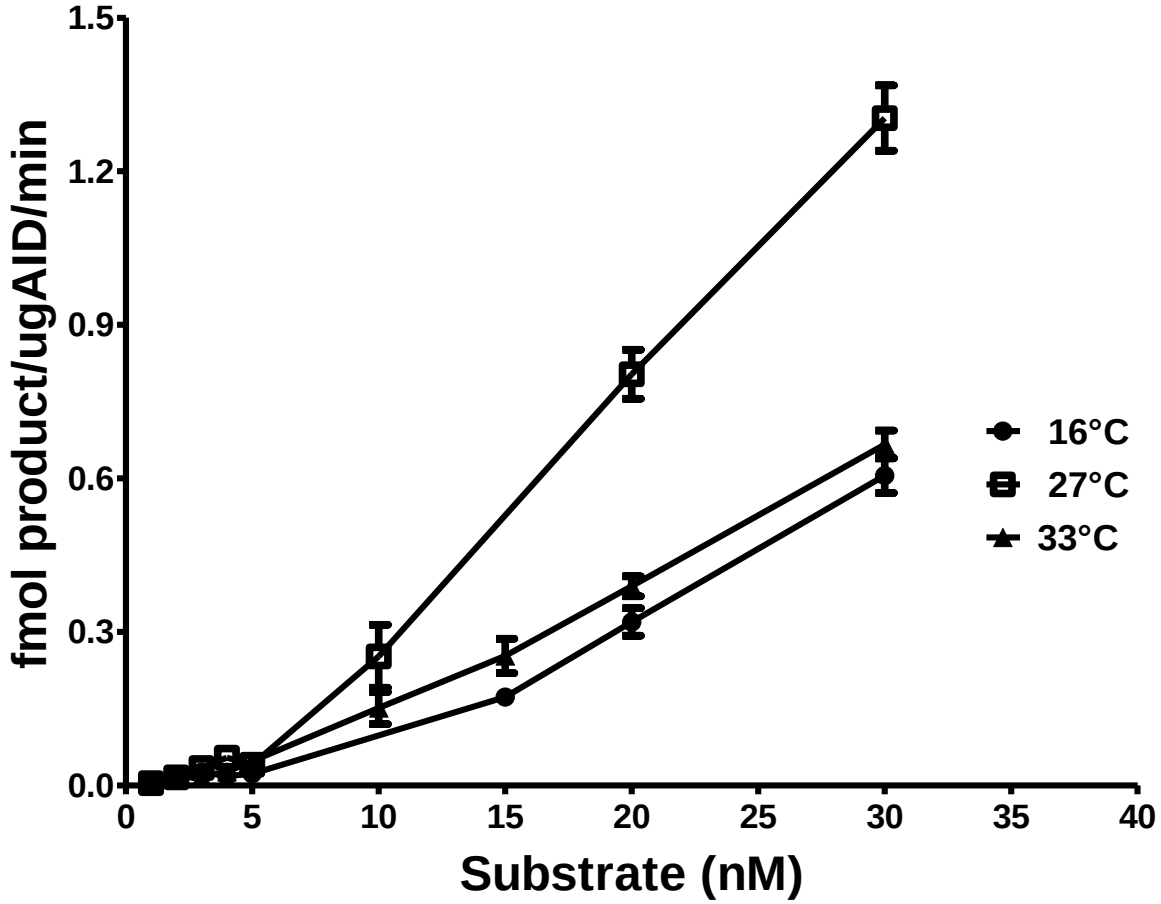


Figure 20: Comparison of deamination kinetics of Dr-AID under different temperature conditions.

Figure 20: Comparison of deamination kinetics of Dr-AID under different temperature conditions.

The deamination kinetics of Dr-AID on TGCbub7 substrate was tested and compared at 16°C, 27°C, and 33°C. The velocity of the reactions was measured as a function of substrate concentration over a 120 fold dilution range (0.25-30 nM). *In vitro*, Dr-AID showed the greatest velocity of product formation at 27°C (the greatest reaction velocities at 16°C, 27°C, and 31°C were 0.60, 1.30, and 0.66 fmol product/ugAID/min, respectively, at the point of maximal difference). The velocity of AID reaction is defined as the amount of deaminated product produced by a given amount of AID enzyme in a unit of time and plotted against substrate concentration.

4. Discussion

4.1 Overall rationale and significant findings

The discovery of AID and its recognition as the master switch that initiates all secondary diversification of antibodies has paved the way for a much deeper understanding of the molecular processes of somatic hypermutation (SHM) and class switch recombination (CSR). AID is expressed in activated mature B lymphocytes and performs its biological function by deaminating cytidines to uridines, thus introducing mutations and/or double-strand breaks throughout the immunoglobulin (Ig) locus. Although that AID activity is mostly directed to the Ig locus, it targets other genes at lower rates and this activity induces genomic lesions leading to transformation of B cells. Despite the fact that numerous studies have been conducted to comprehend the biochemical functions of AID, its structure has not been determined, leaving many questions unanswered about the enzymatic mechanism. Our lab has recently shown that AID from different species have varied deamination activity levels, and that the deamination efficiency of AID from each species is temperature dependent (thermosensitive), such that human AID is more catalytically active at the human body temperature, and less so at lower temperatures (e.g. 25°C & 18°C) (Dancyger et al., 2012). The opposite is true for bony fish AID.

Our goal here was to begin to understand the nature of AID's thermosensitivity by defining the important structures or regions of AID that are responsible for thermosensitivity and by quantitating the degree of differences in activity across temperatures. In particular, we conducted a comparative study to elucidate this property

of human AID versus the most divergent bony fish AIDs. Our aim was to identify region(s) and/or residue(s) within the enzyme which are important for this biochemical characteristic. In doing so, we hoped to elucidate the fundamental relationship between the thermosensitivity of AID and its primary structure, and provide the foundation for doing so with the tertiary structure in the future. Here, we purified human, zebrafish, pufferfish, and, channel catfish AIDs. Hybrid AID enzymes, containing various portions of the aforementioned, were constructed, purified, and tested to narrow down the region(s) that regulate AID thermosensitivity. These hybrid AIDs were tested along with the wild type AID enzymes over a wide range and fine increments of temperatures to precisely determine the optimal temperatures and thermosensitivity patterns of each purified AID enzyme. We hypothesized that the thermosensitivity of AID is a characteristic that is based in one region and can be transferred from one species to another by swapping said region. Otherwise, this property might be attributed to the overall conformational structures of AID in a way that requires residues located throughout the enzyme, to function in combination, rather than a specific region. There are two main findings in this thesis. The first is that the thermosensitivity of AID is mainly a C-terminal based feature. Secondly, we found that human AID reached its maximum enzymatic potential and greatest velocity of product formation at 32°C instead of the previously assumed 37°C.

4.2 The optimal temperature for human AID

Initially, most of our experiments to test human and bony fish AIDs were carried out at three different temperatures 37°C, 25°C, and 18°C. As already rationalized, 37°C was suggested as the optimal temperature for human AID to deaminate its target substrate, based on the facts that the optimum temperature for most human enzymes lies between 35°C and 40°C, and that the body temperature is 37°C. On the other hand, 18°C and 25°C were chosen to represent two distant water temperatures in which bony fish involved in this study naturally live. It is important to note that many fish live in temperature well below this range (e.g. arctic cod), and characterizing their thermosensitivity is a future research direction in our lab. Admittedly, human AID *in vitro* was always more catalytically active at 37°C in comparison to 25°C and 18°C, and its enzymatic activity was diminished as temperature was decreased to such low temperatures (figure 10). One possible explanation for the reduced activity of human AID at 25°C and 18°C could be that human AID at these low temperatures underwent some conformational changes, possibly within the DNA binding groove which would result in reducing the substrate affinity as well as diminishing the catalytic efficiency of the enzyme. In support of this possibility, previous work in our lab showed that fish AID bound DNA with higher affinity at colder temperatures (Dancyger et al., 2012). However, this does not rule out the possibility that a broader conformational change in the enzyme beyond the binding groove is responsible for the diminished activity of human AID. In line with this possibility, another reason for the reduced activity of human AID can be ascribed to low molecular flexibility at colder temperatures. A previous study has revealed that the warm-adapted enzymes in general are structurally rigid and have

reduced catalytic efficiencies at low temperatures (Eijsink, et al., 2004). However, when human AID was tested *in vitro* over wider range and fine increments of temperatures to accurately determine the optimal temperature for human AID, we noticed that human AID is actually optimal at 32°C than at 37°C (figure 12). We confirmed that this optimal activity at 32°C rather than at 37°C, was a true and significant enzymatic property by showing that Hs-AID exhibited greater catalytic velocity at 32°C than at 37°C (figure 19).

This is the first study to show that human AID is optimally active at 32°C. However, if the ideal temperature for human AID *in vivo* is indeed not 37°C, then this *might be* a unique (or uncommon) regulatory feature associated with the biochemical characteristics of this enzyme. In other words, the notion that our body temperature is a slightly sub-optimal temperature for Human AID activity *may* represent an inherently "built-in" mechanism (an evolutionary mechanism) of limiting AID activity, in particular its off-target "deleterious" genome-damaging activities (that cause genomic lesions and lead to transformation of B cells) of Hs-AID in our bodies. In this context, it is interesting to mention that previous work by our group has demonstrated that human AID has a number of unique characteristics such as a very slow rate of deamination activity and unusual high affinity binding to ssDNA (Larijani, et al., 2007) that have also been suggested to restrict its off-target mutagenic activity *in vivo* (Pham et al., 2007; Coker & Petersen-Mahrt, 2007; Larijani & Martin, 2012; Dancyger et al., 2012). Alternatively, this evolutionary feature may not serve an important biological function, but may simply be a by-product of the evolution of AID to become more warm-adapted than fish AID, but to a level that has not yet reached 37°C.

4.3 Thermosensitivity mechanisms of human and bony fish AIDs

Earlier studies have shown that various species have evolved in dissimilar conditions, causing them to adapt to different thermal environments (Feller & Gerday, 2003; Smalas, 2000; D'Amico et al., 2002; Somero, 2004; Siddiqui & Cavicchioli, 2006). Also, previous work in our lab has shown that human and bony fish AIDs have disparate enzymatic activities when they are subjected to different thermal conditions (Dancyger et al., 2012). Here we showed that human AID *in vitro* has unexpectedly obtained its maximum deamination activity at 32°C (figure 12), but it was catalytically sensitive to temperatures lower than 32°C. In addition, some recent lines of evidence indicate that warm-adapted enzymes (e.g. human enzymes) are generally characterized by being structurally rigid at low temperatures (Eijsink, et al., 2004). This may be the reason for the reduced deamination activity of human AID observed at temperatures lower (or even higher) than 32°C. In contrast, we noticed that bony fish AIDs involved in our study are adapted to temperatures around 25°C, but their catalytic rates decreased as temperature was increased (figure 12). This was not surprising, since these are natural water temperatures where some bony fish (e.g., zebrafish and channel catfish) normally live (Shrable et al., 1969; Chen 1976; Spence et al., 2006). These findings brought forth the question of what kind of adjustments have resulted in the adaptation of human and bony fish AIDs to disparate temperatures. One possible explanation for the cold adaptation of bony fish AIDs could be that these enzymes may have developed some structural features (e.g. divergent amino acid residue(s) or different loop structure(s) at the primary structural level or at the three dimensional level, respectively, that somehow provided these enzymes with a higher molecular flexibility, particularly around the catalytic site,

and high specific deamination activity at low temperatures compared to the warm-adapted enzymes. In this context, it is noteworthy to mention that some studies showed that the non-catalytic regions (N-terminal and C-terminal regions) of the cold-adapted enzymes have preserved relative stability, while the catalytic regions have evolved certain plasticity to bind the substrate at low temperatures (D'Amico S, et al., 2002). An interesting finding in our study was that human AIDs whose N-terminal or catalytic regions were interchanged with the equivalent regions of bony fish AIDs (e.g. HDH, HIH, DHH, THH, DDH) did not alter their thermosensitivity patterns and their optimal temperatures remained very close to that of human AID (figure 11-A, 14, 16, and 18). This suggests that swapping these regions (N-terminal or catalytic regions) between human and bony fish AIDs does not transfer the optimal temperatures or the thermosensitivity from one species to another. In good agreement with our findings, a previous study on a hybrid human AID (composed of human AID with a catalytic region of takifugu AID), using both a bacterial rifampicin-rescue assay and a GFP-stop-codon-reversion assay done in HEK293 cells, revealed that the thermosensitivity of this hybrid human AID was not affected and still has strong enzymatic activity at 37°C (Ichikawa et al., 2006; Barreto & Magor, 2011). These results further confirmed that this biochemical property of AID lies outside the catalytic (Zn coordinating) region, and it also provided additional *in vivo* relevance to our observations, by showing that the same trends we are measuring in our *in vitro* assays on purified GST-AID have been observed when studying AID activity inside cells.

A previous study on the three dimensional structures of cold-adapted enzymes revealed that the amino acid residues located in the active site of these enzymes are

generally highly conserved with their warm-adapted equivalents, and only few conformational adjustments/modifications are required for the warm-adapted enzymes to adapt to low temperatures, and that these minor modifications responsible for the adaptive changes have to be located outside the active site (Field & Somero, 1998; D'Amico et al., 2002). All our findings thus far support the notion that the thermosensitivity of AID is mostly a C-terminal region-based characteristic and swapping this region can transfer this feature from one species to another (figure 18). However, it remains to be determined whether this may be a single amino acid residue that harbors this biochemical property or a combination of amino acids or small segments that might be continuously or dis-continuously located in the C-terminal region of AID.

4.4 Conclusion and future directions

In this study, we endeavoured to shed light on the temperature-adaptive variation between human and bony fish AIDs. For that reason, AID of human, zebrafish, catfish and pufferfish were purified and examined over a wide range of temperatures. Intriguingly, human AID was optimally active at 32°C and not at 37°C as we initially expected, whereas the average ideal temperatures for bony fish AIDs examined in this study were 22°C for catfish AID and 27°C for zebrafish AID. Also, we compared the deamination kinetics of human AID at 37°C versus 32°C; this further confirmed that human AID is more catalytically robust at 32°C. We also sought to determine the portions of AID's primary structure that led to these adaptive differences in the thermosensitivity of AID among species. Hybrid AID enzymes composed of various segments of human and bony fish AIDs were constructed, purified, and tested along with the wild type AIDs under the same substrate and temperature conditions. We found that the thermosensitivity of AID is mediated by the C-terminal region. However, since our hybrid AIDs were constructed based on dividing the primary structure of AID into three linear regions, and not based on structure modeling, more ongoing and future work is required to design new hybrids by swapping regions that encode putative three-dimensional structural features. Extensive modeling will be used to identify divergent structural loops in the C-terminal region of human and bony fish AIDs. The loops that have divergent conformational structures will be swapped between species with the premise that this may influence the thermosensitivity and reveal "structural features" rather than primary sequences (as was done here) that regulate this enzymatic feature of AID. Also, more species (e.g. various cold and warm adapted organisms) will be included in future studies on this biochemical

property of AID. It will be of particular interest to test the thermosensitivity of AID from bony fish that inhabit extremely cold temperatures such as the North Atlantic and Arctic. In addition, specific mutants of AID, focusing on the C terminus region, will be generated harboring multiple mutations that have altered their thermosensitivity; in order to more precisely map the origin and mechanism of this enzymatic characteristic. Furthermore, circular dichroism (CD) can be utilized to study and compare the folding mechanism and stability of the cold- *vs.* warm-adapted AIDs at different temperatures, which will test the "unfolding" hypothesis presented in this thesis. In the long-term, it will also be of interest to determine the mutagenic activity of human AID at different temperatures (e.g. at 32°C *vs*. 37°C) by measuring double stranded breaks mediated by this enzyme inside cells, to further confirm the notion that 37°C is a slightly sub-optimal temperature for Hs-AID activity and that this may represent an inherently built in restriction to AID's off-target activities.

5. References

Abdouni H, King J, Suliman M, Quinlan M, Fifield H, Larijani M., 2013. Zebrafish AID is capable of deaminating methylated deoxycytidines. *Nucleic Acids Research*, 1(14), 5457-68.

Barreto VM, Magor BG., 2011. Activation-induced cytidine deaminase structure and functions: A species comparative view. *Developmental & Comparative Immunology*, 35(9), 991-1007.

Beale RC, Petersen-Mahrt SK, Watt IN, Harris RS, Rada C, Neuberger MS., 2004. Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo. *J. Mol. Biol*, 337, 585–596.

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

Ronda Bransteitter, Phuong Pham, Matthew D, Scharff 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(7), 4102-7.

Chiu Y, Greene W., 2008. The APOBEC3 cytidine deaminases: an innate defensive network opposing exogenous retroviruses and endogenous retroelements. *Annu Rev Immunol*, 26, 317-53.

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

Silvestro G Conticello., 2008. The AID/APOBEC family of nucleic acid mutators. *Genome Biology*, 6, 9-229.

Silvestro G Conticello, Thomas Svend, K Petersen-Mahrt., 2005. Evolution of the AID/APOBEC Family of Polynucleotide (Deoxy)cytidine Deaminases. *Mol Biol Evol*, 22(2), 367-77.

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

Salvino D'Amico, Paule Claverie, Tony Collins, Daphné Georlette, Emmanuelle Gratia, Anne Hoyoux, Marie-Alice Meuwis, Georges Feller, Charles Gerday., 2002. Molecular basis of cold adaption. *Philos. Trans. R. Soc. London Ser. B*, 357, 917–25.

Dancyger A, King J, Quinlan M, Fifield H, Tucker S, Saunders H, Berru M, Magor B, Martin A, Larijani M., 2012. Differences in the enzymatic efficiency of bony fish and human AID are mediated by a single residue in the C-terminus that modulates single-stranded DNA binding. *FASEB J*, 26(4), 1517-25.

deYébenes V, Ramiro A., 2006. Activation-induced deaminase: light and dark sides. *Trends Mol Med*, 12, 432-9.

Dickerson S K, Eleonora M, Besmer E, Papavasiliou N., 2003. AID Mediates Hypermutation by Deaminating Single Stranded DNA. *J Exp Med*, 197(10), 1291-1296.

Durandy A, Peron S, Fischer A., 2006. Hyper-IgM syndromes. *Curr Opin Rheumatol*, 18(4), 369-76.

Eijsink VG, Bjørk A, Gåseidnes S, Sirevåg R, Synstad B, van den Burg B, Vriend G., 2004. Rational engineering of enzyme stability. *J. Biotechnol*, 113, 105–20.

Feller G, Gerday C., 2003. Psychrophilic enzymes: hot topics in cold adaptation. *Nat Rev Microbiol*, 1(3), 200-8.

Fields PA., 2001. Protein function at thermal extremes: balancing stability and flexibility. *Comp Biochem Physiol A Mol Integr Physiol*, 129, 417-31.

Fields P, Houseman D., 2004. Decreases in activation energy and substrate affinity in cold-adapted A4-lactate dehydrogenase: evidence from the Antarctic notothenioid fish Chaenocephalus aceratus. *Mol Biol Evol*, 21(12), 2246-55.

Fields P, Somero G., 1998. Hot spots in cold adaptation: Localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. *Proc. Natl. Acad. Sci. USA*, 95, 11476–81.

Fleisher T A, Tomar R H., 1997. Introduction to Diagnostic Laboratory Immunology. *JAMA: The Journal of the American Medical Association*, 278(22), 1823-34

Gellert M., 2002. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem*, 71, 101-32.

Georlette D, Blaise V, Collins T, D'Amico S, Gratia E, Hoyoux A, Marx JC, Sonan G, Feller G, Gerday C., 2004. Some like it cold: biocatalysis at low temperatures. *FEMS Microbiol Rev*, 28(1), 25-42.

Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M, Honjo T., 2004. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like

apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc Natl Acad Sci U S A,* 101(7), 1975-80.

Jankovic M, Robbiani DF, Dorsett Y, Eisenreich T, Xu Y, Tarakhovsky A, Nussenzweig A, Nussenzweig MC., 2010. Role of the translocation partner in protection against AID-dependent chromosomal translocations. *Proc.Natl. Acad. Sci. U.S.A,* 107, 187-192.

Kinoshita K, Honjo T., 2001. Linking class-switch recombination with somatic hypermutation. *Nat Rev Mol Cell Biol,* 2(7), 493-503.

Komori J, Marusawa H, Machimoto T, Endo Y, Kinoshita K, Kou T, Haga H, Ikai I, Uemoto S, Chiba T., 2008. Activation-induced cytidine deaminase links bile duct inflammation to human cholangiocarcinoma. *Hepatology,* 47, 888-896..

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

Lanasa M C, Weinberg B., 2011. Immnuophenotypic and Gene Expression Analysis of Monoclonal B Cell Lymphocytosis Shows Biologic Characteristics Associated With Good Prognosis CLL. *Leukemia,* 25(9), 1459-66.

Larijani M, Frieder D, Basit W, Martin A., 2005. The mutation spectrum of purified AID is similar to the mutability index in Ramos cells and in ung(-/-)msh2-/- mice. *Immunogenetics,* 56, 840-5.

Larijani M, Martin A., 2007. Single-stranded DNA structure and positional context of the target cytidine determine the enzymatic efficiency of AID. *Mol. Cell. Biol,* 27, 8038–8048.

Larijani M, Martin A., 2012. The Biochemistry of Activation-Induced Deaminase and its physiological functions. *Seminars in Immunology,* 24(4), 255-63.

Larijani M, Petrov A P, Kolenchenko K., 2007. AID associates with single-stranded DNA with high affinity and a long complex half-life in a sequence-independent manner. *Mol. Cell. Biol,* 27, 20-30.

Litman G W, Cannon J P, Dishaw L., 2005. Reconstructing immune phylogeny: new perspectives. *Nature Reviews Immunology,* 5(11), 866–79.

Li Z, Caroline J, Woo M D, Iglesias-Ussel S., 2004. The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes & Dev,* 18, 1-11.

Losey H, AJ R, GL V., 2006. Crystal structure of Staphylococcus aureus tRNA adenosine deaminase TadA in complex with RNA. *Nat Struct Mol Biol,* 13, 153-159.

Martin A, Scharff M., 2002. AID and mismatch repair in antibody diversification. *Nature Reviews Immunology,* 2, 605-614.

Matsuda F, Ishii K, Bourvagnet P, Kuma Ki, Hayashida H, Miyata T, Honjo T., 1998. The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus. *The Journal of experimental medicine ,* 188(11), 2151–62.

Mechtcheriakova D, Svoboda M, Meshcheryakova A, Jensen-Jarolim E., 2012. Activation-induced cytidine deaminase (AID) linking immunity, chronic inflammation, and cancer. *Cancer Immunol. Immunother,* 61, 1591-1598.

Morisawa T, Marusawa H, Ueda Y, Iwai A, Okazaki IM, Honjo T, Chiba T., 2008. Organ-specific profiles of genetic changes in cancers caused by activation-induced cytidine deaminase expression. *Int. J. Cancer,* 123, 2735-2740.

Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, Honjo T., 1999. Specific Expression of Activation-induced Cytidine Deaminase (AID), a Novel Member of the RNA-editing Deaminase Family in Germinal Center B Cells. *J. Biol. Chem,* 274, 18470-18476.

Navaratnam N, Morrison JR, Bhattacharya S, Patel D, Funahashi T, Giannoni F, Teng BB, Davidson NO, Scott J., 1993a. The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. *J. Biol. Chem,* 268, 20709–20712.

Neuberger MS, Di Noia JM, Beale RC, Williams GT, Yang Z, Rada C., 2005. Somatic hypermutation at A·T pairs: polymerase error versus dUTP incorporation. *Nat. Rev. Immunol,* 5, 171-78.

Neuberger MS, Ehrenstein MR, Rada C, Sale J, Batista FD, Williams G, Milstein C., 2000. Memory in the B-cell compartment: antibody affinity maturation. *Philos Trans R Soc Lond B Biol Sci,* 355(1395), 357-60.

Odegard V, DG S., 2006. Targeting of somatic hypermutation. *Nat Rev Immunol,* 6(8), 573-83.

Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K, Honjo T., 2003. Constitutive Expression of AID Leads to Tumorigenesis. *JEM,* 197(9), 1173-1181 .

Or-Guil M, Wittenbrink N, Weiser A, Schuchhardt J., 2007. Recirculation of germinal center B cells: a multilevel selection strategy for antibody maturation. *Immunol Rev,* 216, 130–41.

Pasqualucci L, Neumeister P, Goossens T, Nanjangud G, Chaganti RS, Küppers R, Dalla-Favera R., 2001. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature,* 412(6844), 341-6.

Jonathan U Peled, Fei Li Kuang, Maria D Iglesias-Ussel, Sergio Roa, Susan L Kalis, Myron F Goodman, Matthew D Scharff., 2008. The biochemistry of somatic hypermutation. *Annu. Rev. Immuno,* 26, 481–511.

Pham P, Chelico L, Goodman M F., 2007. DNA deaminases AID and APOBEC3G act processively on single-stranded DNA. *DNA Repair,* 6, 689–692.

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

Prochnow C, Bransteitter R, MG K, Goodman MF., 2007. The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature,* 445, 447-451.

Rajewsky K., 1996. Clonal selection and learning in the antibody system. *Nature,* 381, 751-758.

Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R, Gennery A, Tezcan I, Ersoy F, Kayserili H, Ugazio AG, Brousse N, Muramatsu M, Notarangelo LD, Kinoshita K, Honjo T, Fischer A, Durandy A., 2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell,* 102(5), 565-75.

Saunders H L, Magor B G., 2004. Cloning and expression of the AID gene in the channel catfish. *Dev. Comp. Immunol,* 28, 657–663.

Saunders HL, Oko AL, Scott AN, Fan CW, Magor BG., 2010. The cellular context of AID expressing cells in fish lymphoid tissues. *Developmental and Comparative Immunology,* 34, 669-676.

Seok-Rae Park., 2012. Activation-induced Cytidine Deaminase in B Cell Immunity and Cancers. *Immune Netw,* 12(6), 230–239.

Sheehy A, Gaddis N, Choi J, Malim M., 2002. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature,* 418(6898), 646-50.

Siddiqui K S, Cavicchioli R., 2006. Cold-Adapted Enzymes. *Annu. Rev. Biochem*, 75, 403–33.

Sohail A, Joanna K, Mala S, Asad Ullah., 2003. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Res,* 31(12) 2990–2994.

Stavnezer J, Guikema JE, Schrader CE., 2008. Mechanism and regulation of class switch recombination. *Annu Rev Immunol,* 26, 261-92.

B Teng, CF Burant, NO Davidson., 1993. Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science,* 260(5115), 1816-9.

Warrington R, Watson W, Kim HL, Antonetti FR., 2011. An introduction to immunology and immunopathology. *Allergy Asthma Clin Immunol,* 10(7), 1492-7.

Woof J M, Burton D R., 2004. Human antibody–Fc receptor interactions illuminated by crystal structures. *Nature Reviews Immunology,* 4(2), 89–99.

Yang S, Schatz D., 2007. Targeting of AID-mediated sequence diversification by cis-acting determinants. *Adv Immunol,* 94, 109-125.

Yu Q, König R, Pillai S, Chiles K, Kearney M, Palmer S, Richman D, Coffin JM, Landau NR., 2004. Single-strand specificity of APOBEC3G accounts for minus-strand deamination of the HIV genome. *Nat Struct Mol Biol,* 11(5), 435-42.

Zavodszky P, Kardos J, Svingor Á, Petsko A., 1998. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *PNAS,* 95(13), 7406-7411.

Zhao Y, Pan-Hammarström Q, Zhao Z, Hammarström L., 2005. Identification of the activation-induced cytidine deaminase gene from zebrafish: an evolutionary analysis. *Devel. Comp. Immunol,* 29, 61-71.

Ziqiang Li, Caroline J Woo, Maria D Iglesias-Ussel, Diana Ronai, and Matthew D Scharff., 2004. The generation of antibody diversity through somatic hypermutation and class switch recombination. *Genes Dev,* 18, 1-11.