Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Immunology

AID hits the jackpot when missing the target

Amy L Kenter, Satyendra Kumar, Robert Wuerffel and Fernando Grigera

Activation induced deaminase is the single B cell specific factor mediating class switch recombination and somatic hypermutation. Numerous studies have shown that AID preferentially targets Ig substrates and also attacks non-Ig substrates to create DNA damage that contributes to lymphomagenesis. AID targeting to Ig loci is linked to transcription but the mechanism governing this process has been obscure. Here we discuss research that illustrates the connection between AID targeting to DNA substrates and transcription processes to reveal rules governing the specificity of AID attack. These observations are woven together to provide a integrated view of AID function and a surprising linkage with global regulation of gene expression.

Address
Department of Microbiology and Immunology, University of Illinois College of Medicine, Chicago, IL, USA

Corresponding author: Kenter, Amy L (star1@uic.edu)

Current Opinion in Immunology 2016, **39**:96–102

This review comes from a themed issue on **Lymphocyte development and activation**

Edited by **James Hagman**

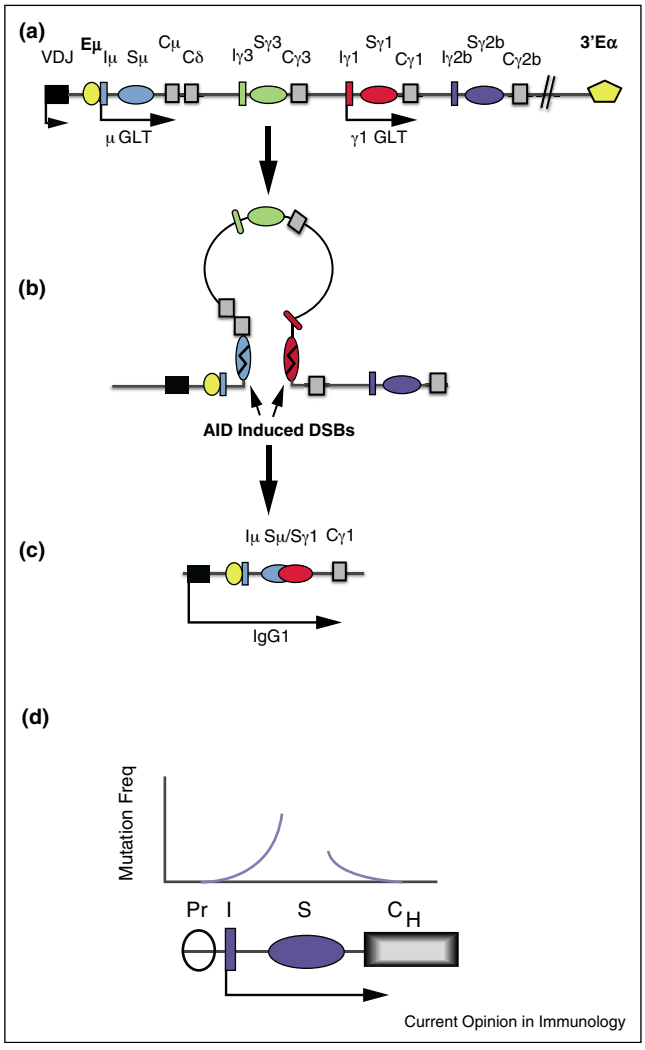
http://dx.doi.org/10.1016/j.coi.2016.01.008

Humoral immunity is mediated in B cells by antigen receptors (BCR) that are composed of immunoglobulin (Ig) heavy (H) and light (L) chains. The antigen receptor component of Ig is assembled from multiple gene segments via V(D)J recombination during early B cell development in the bone marrow and is mediated by the RAG1 and RAG2 recombinases [1]. In the peripheral lymphoid organs, B cells become activated by antigen and undergo somatic hypermutation (SHM) and class switch recombination (CSR) and upon terminal differentiation secrete Ig as antibody [2,3]. Antibody repertoires are diversified at high frequency during somatic hypermutation (SHM), by introducing mutations into *Igh* and *Igl* V(D)J exons. SHM occurs in germinal centers of the peripheral lymphoid organs where mutations are selected to produce BCR with increased affinity. Constant (C_H) region genes encode the C-terminal domains of the Ig chains and determine IgH effector activities. IgH effector function is diversified through CSR, while retaining the original antigen binding specificity arising during V(D)J recombination. Activation induced deaminase (AID), a cytosine deaminase is essential for initiating both SHM and (CSR) in mature B cells [4••].

The mouse *Igh* locus includes eight C_H genes, encoding μ, δ, γ3, γ1, γ2b, γ2a, ε, and α chains and each is paired with repetitive switch (S) DNA (with the exception of Cδ). The C_H region genomic area spans 220 kb and is flanked by the intronic Eμ and 3′Eα enhancers [5]. CSR is focused on S regions and involves intra-chromosomal deletional rearrangements that replace the initial Cμ with a downstream C_H region gene (Figure 1). Prior to and during CSR, a S-S synaptosome is formed to facilitate proximity between the donor Sμ and a downstream acceptor S region [6••,7–9]. Formation of the S-S synaptosome is critically dependent on long range chromatin interactions that are tethered by key transcriptional elements (reviewed in [10]). AID initiates CSR by deaminating cytosines in S regions in donor Sμ and a downstream acceptor S region (Box 1) [11]. AID dependent DNA damage is focused to highly degenerate WRC hotspot motifs (W = A/T, R = A/G) that are found in V genes and at high density in S regions (reviewed in [12]). Conversion of AID induced lesions to DNA double strand breaks by general DNA repair factors has been extensively reviewed [2,3,11–17]. Degeneracy of the AID hotspot motif raises questions regarding the mechanism by which AID is targeted to its Ig substrates.

Transcription is a hallmark of SHM and CSR and transcriptional elements are candidates for providing specificity to AID targeting [5,18]. V region transcription initiates from a promoter 5′ proximal to the rearranged V(D)J exon and terminates 3′ of the C_H region exons. CSR is focused to specific S regions by differential activation of germline transcription [10,19]. C_H gene transcription units are comprised of the noncoding intervening (I) exon, an S region and a C_H coding region. Germline transcription initiates at a transcription start site (TSS) 5′ of each I exon, proceeds through the S region and terminates downstream of the corresponding C_H gene (Figure 1a). The V(D)J mutation profile has a sharp 5′ boundary ~120 bp downstream of the TSS and a less defined 3′ border ~1 kb downstream of the promoter. Alteration of the V region promoter position displaces transcription initiation and perturbs the mutation distribution [20,21] thereby linking induction of AID

Figure 1



The looping-out and deletion model of Ig switch recombination. **(a)** A partial schematic map of *Igh* locus before CSR is shown not to scale. A productive V(D)J rearrangement has occurred allowing expression of the γ and ^{TM} IgH chains. Intact Sμ and Sγ1 are separated by approximately 70 kb. Stimulation of B cells with antigen or mitogen induces germline transcription through the Iγ1-Sγ1-Cγ1 region before recombination. **(b)** The Sμ and Sγ1 regions are aligned causing the intervening genomic DNA to form a loop. **(c)** A reciprocal crossover between Sμ and S©1 results in the formation of a new hybrid transcriptional unit containing the original VDJ exons contiguous with Cγ1 and the formation of hybrid Sμ/S©1 molecules. **(d)** A schematic diagram depicts a generic I-S-C_H region downstream of the GLT promoter (Pr). Above the diagram a summary of mutation frequencies 5′ and 3′ of the S region is shown.

dependent mutations with transcription. Similarly, AID induced DNA lesions in S regions begin ~150 bp downstream of the I exon TSS [22]. A recent study shows that sequence intrinsic features target AID dependent DNA lesions to Ig templates [23]. Although AID targeting to Ig substrates requires transcription, the unique transcriptional features that determine preferential AID attack at Ig templates have been difficult to discern.

Box 1 AID introduces DNA damage by converting deoxycytidine (dC) to deoxyuracil (dU) [14]. The AID initiated DNA lesions are processed by engagement with base excision repair (BER) and mismatch repair (MMR) pathways to create mutations required for SHM and DNA double strand break (DSB) intermediates that consumed in CSR (reviewed in [2,3,11,14]). The observations that AID dependent dU residues are detected in the 5′Sμ region [99] and that uracil DNA glycosylase (UNG) is required for formation of double strand breaks (DSB) and mutations demonstrate that AID deamination initiates SHM and CSR [100,101]. Conversion of AID induced lesions to DNA double strand breaks is mediated by general DNA repair factors [17].

AID induced double strand breaks (DSBs) in normal B cells occur at hundreds of non-Ig sites many of which are syntenic with sites of translocations, deletions, and amplifications found in human B cell lymphomas [24]. Physiological levels of AID in GC B cells have been linked to deamination of a large cohort of non-Ig genes [25–28] where the mutation rate is 20–100 fold lower than at Ig loci [26]. The findings that SHM and CSR are tightly linked to transcription [20,29,30] and AID interacts with RNA Pol II (RNAP II) [27,31] have led to the notion that transcription and RNAP II-associated proteins might facilitate the binding of AID to target DNA sequences. Based on provocative new studies, we address three important interrelated questions regarding AID targeting. What is the mechanism by which AID preferentially targets Ig substrates and its corollary, how is AID attack at non-Ig substrates directed? Finally, is there an adaptive advantage for AID attack at non-Ig templates? We focus primarily on CSR as this area has been the most intensively investigated. Recent work implies that AID attack on non-Ig substrates may mediate functional outcomes that contribute to a biological 'jackpot'.

Transcribed S regions are specialized targets of AID

S regions are repetitive, nonidentical, 1–12 kb long and guanine rich on the nontemplate strand [32]. Deletion, inversion or replacement of S regions reduces CSR frequency indicating that S regions are specialized targets of CSR [33–36]. However, CSR S/S junctions are notable for their lack of consensus sequence or homology originating from site specific or homologous recombination. The degeneracy of the S region repeats and the absence of discernable recombination signal motifs have led to models in which higher order structures provide recognition motifs for the CSR machinery. *In vitro* transcription studies indicated that the looped out ssDNA nontemplate strand can assume specialized structures including stem loops [37], four-stranded G quartets [38] and R-loops (Box 2) [39,40]. *In vivo* studies confirm that transcription through S regions *in vivo* creates G quartets [41,42] and