Seminars in Immunology 24 (2012) 293–300

Contents lists available at SciVerse ScienceDirect

Seminars in Immunology

journal homepage: www.elsevier.com/locate/ysmim

Review

AIDing antibody diversity by error-prone mismatch repair

Richard Chahwan[a], Winfried Edelmann[a], Matthew D. Scharff[a,*], Sergio Roa[b,**]

[a] Department of Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA
[b] Oncology Division, Center for Applied Medical Research (CIMA), University of Navarra, 31008 Pamplona, Spain

ARTICLE INFO

Keywords:
Mismatch repair
Class switch recombination
Somatic hypermutation
Activation-induced deaminase
Double-strand breaks
Cytosine deamination
Epigenetic
Antibody diversity

ABSTRACT

The creation of a highly diverse antibody repertoire requires the synergistic activity of a DNA mutator, known as activation-induced deaminase (AID), coupled with an error-prone repair process that recognizes the DNA mismatch catalyzed by AID. Instead of facilitating the canonical error-free response, which generally occurs throughout the genome, DNA mismatch repair (MMR) participates in an error-prone repair mode that promotes A:T mutagenesis and double-strand breaks at the immunoglobulin (Ig) genes. As such, MMR is capable of compounding the mutation frequency of AID activity as well as broadening the spectrum of base mutations; thereby increasing the efficiency of antibody maturation. We here review the current understanding of this MMR-mediated process and describe how the MMR signaling cascade downstream of AID diverges in a locus dependent manner and even within the Ig locus itself to differentially promote somatic hypermutation (SHM) and class switch recombination (CSR) in B cells.

1. Introduction

Organisms at every level of evolutionary development are constantly exposed to genotoxic stress that can damage and alter the structure of their genetic material. To counteract this threat, a large variety of highly conserved repair mechanisms have been generated throughout evolution [1]. It is estimated that 2–4% of the genes in mammalian cells are devoted to the repair of DNA damage and these are tightly integrated through various signal transduction pathways with the regulation of the cell cycle and cell death [2]. Paradoxically, the generation and propagation of "useful" mutations could potentially confer an evolutionary advantage [3]. Furthermore, in complex organisms it is sometimes important to somatically produce increased levels of genetic diversity [4,5].

Perhaps the most extreme example of engineering and tightly regulating genomic instability for a selective advantage is the adaptive immune response in higher organisms, where it is essential to spawn an enormous repertoire of antigen binding sites in B and T cell antigen receptors. This is achieved by the combinatorial rearrangement of a small number of variable (V), diversity (D) and joining (J) genes so that the organism can mount a protective immune response against most foreign antigens that it encounters [6]. B cells need to generate antibodies of high affinity to neutralize and inactivate pathogenic agents in the blood stream, in tissues and even in mucosal spaces. To achieve this V(D)J regions that encode low affinity antibodies are somatically hypermutated (SHM) so that they achieve the high affinities required to neutralize toxic foreign agents [7,8]. In order to distribute these protective antibodies throughout the body and enable them to carry out different effector functions, it is also necessary to mutate the switch (S) regions that are upstream from each of the constant region genes and to convert those mutations into double stranded DNA breaks (DSBs). This allows the heavy chain V(D)J regions encoding the antigen-binding site to be rearranged from the μ constant region to the downstream γ, ε, and α constant regions through a process, termed class switch recombination (CSR) [9,10].

In B cells, a potent mutator known as activation-induced deaminase (AID), initiates SHM of the Ig V(D)J and CSR of the Ig S regions by deaminating Cs and generating U:G DNA mismatches at a very high frequency. This occurs primarily in the activated B cells in the germinal centers of secondary lymphoid organs, such as the lymph nodes, spleen and tonsils where AID is expressed at very high levels. In those germinal center B cells, AID induced mutations recruit base excision (BER) and mismatch repair (MMR) processes that in other cell types and at non-Ig genes repair DNA lesions with high fidelity [11,12]. However, in B cells, the BER and MMR that are recruited by AID-induced U:G mismatches facilitate an error-prone repair of these mismatches, and MMR is responsible for as many as half of the mutations that arise during SHM and for most of the mutations that occur at A:T bases [13]. While a great deal is known from biochemical and yeast studies about the role of MMR proteins in the normal

* Corresponding author at: Department of Cell Biology, The Albert Einstein College of Medicine, 1300 Morris Park Ave., Chanin 404, Bronx, NY 10461, USA. Tel.: +1 718 430 3527; fax: +1 718 430 8574.
** Corresponding author at: Center for Applied Medical Research (CIMA), University of Navarra, 31008 Pamplona, Spain. Tel.: +34 948 194700x1028; fax: +34 948 194714.
E-mail addresses: matthew.scharff@einstein.yu.edu (M.D. Scharff), sroa@unav.es (S. Roa).

1044-5323/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.smim.2012.05.005

process of high fidelity MMR, less is known about how these factors are targeted to particular mismatches *in vivo* and to the regulation of the repair process in different lineages of mammalian cells.

In this review we will describe the orchestration of MMR-mediated error-prone repair in both antibody V(D)J and S regions following the enzymatic production of DNA mismatches by AID. We will also point out the many unresolved features of this atypical and potentially dangerous repair process. Through comparison of the complete loss and separation-of-function mutants in the MMR proteins *in vivo* and the detailed examination of the outcome of this process as it is reflected in the sequences of individual antibody V(D)J and S regions, we are learning new aspects of how error-prone MMR generates antibody diversity in B cells. We are also gaining new insights into how high fidelity MMR is regulated in general and how its misregulation can lead to tumorigenesis.

2. AID-mediated cytosine deamination instigates a highly mutagenic cascade

Because it is extremely mutagenic, highly expressed in centroblast B cells, and generates mutations characteristic of antibody V and S regions, it was originally thought that AID was a B cell specific deaminase primarily devoted to the generation of antibody diversity [8]. However, subsequent studies revealed that many other genes were mutated in activated B cells some of which were repaired with high fidelity while others were also subjected to error-prone repair [14,15] (see Saribasak and Gearhart, in this issue). The recent report that there are ~1 million sites occupied by AID in activated mouse B cells is surprising considering its genotoxic potential [16]. This apparent promiscuity of targeting may, to some extent, be explained by the fact that AID mediated cytosine deamination can also cause the active demethylation of 5-methylcytosines or 5-hydroxymethylcytosines with potential consequences for protein expression and epigenetic inheritance. The finding of AID in germ cells and the demonstration of a role for AID-induced active demethylation in early differentiation and even in neuronal function [17–20] may, to some extent, explain why AID is so widely distributed. However, in B cells AID seems to cause mutations to various extents in most of the genes that it targets [15,16].

AID selectively deaminates the C in WRCY (W = A/T, R = A/G, Y = C/T) generating a U:G mismatch (WRUY). While this motif is enriched in the parts of the V(D)J regions that form the antigen-binding site and in the S regions, this alone does not explain the specific targeting to the Ig locus since the WRCY motif is widely distributed throughout the genome. The finding that single-stranded DNA (ssDNA) is the substrate for AID in part explains why high rates of transcription, accompanied by the high potential for transcription stalling [21], are required for AID mutagenesis. Indeed, ssDNA accessible regions and transcription bubbles provide an excellent substrate for AID [22–24], but this does not explain why only some but not all highly transcribed genes are targeted by AID in B cells [16]. As will be discussed in other chapters in this issue (Kenter, in this issue), some cis-acting sequences such as E2A binding sites [25,26] and chromatin modifications provide a partial explanation of how AID mutagenesis might be largely restricted to the V(D)J and S regions of the Ig gene [27–30], but these are also very widely distributed throughout the genome. Therefore, the question remains whether there is any specific genetic or epigenetic signature that licenses AID and error-prone repair to target the V and S regions of the Ig locus more frequently and efficiently than other parts of the Ig locus and non-Ig genes.

Mutations in AID that lead to a loss of SHM and CSR in patients with Hyper-IgM Syndrome type II and subsequent site directed mutagenesis and exon swapping experiments suggested that there are likely to be associated proteins and post-translational modifications that regulate AID levels and its concentration in the nucleus, and also selectively target AID to parts of the Ig gene. While a number of associated proteins have been discovered (see Orthwein and Di Noia, in this issue; Larijani and Martin, in this issue; Häsler et al., in this issue), some of which like RPA and Spt5 probably maintain the single stranded nature of V(D)J and S regions [21], it is still unclear how much of the rest of the genome is protected from the mutagenic AID activity [15]. These factors also do not explain how AID is regulated to carry out active demethylation as opposed to hypermutation [31]. In fact, it has been suggested that a different set of associated proteins including Gadd45α and glycosylases like MBD4 and TDG are required for AID to act as a promoter of cytosine demethylation through its deaminase activity [32,33]. These issues have become even more important as it is now clear that AID is expressed, at least at low levels, in cells of many lineages, and that its mutagenic activity may not only be responsible for B cell malignancies, but also for many other types of cancers [34,35]. It has thus become very important to understand exactly how AID is recruited to the Ig V(D)J and S regions. It is possible that the factors or DNA and chromatin structures that are responsible for the selective targeting of AID may also cause MMR to be error-prone at the Ig gene and in some of the other genes that are targeted by AID [15].

3. MMR mediates the resection of ssDNA patches and the introduction of A:T mutations

The current model (Fig. 1) suggests, the U:G mismatch that is generated when AID mutates an Ig V(D)J region is either replicated over to produce a U → T mutation, recognized by UNG to initiate a sometimes error-prone BER (see also Saribasak and Gearhart, in this issue), or recognized by the MutSα MMR heterodimer composed of the MSH2 and MSH6 proteins (Fig. 1). This MMR-mediated sensing of the U:G lesion initiates a series of processes that are responsible for the mutations at surrounding A:T bases. When the U:G mismatch is recognized by MutSα, ATP-mediated conformational changes [36–38] allow it to recruit PCNA and 5′–3′ exonucleases, such as EXO1 [39]. The mismatch is subsequently excised to create a single stranded patch. The excised strand is then replaced by a new strand of DNA that might acquire additional mutations at A:T bases. This error-prone resynthesis is mediated, at least in part, by the polymerase eta (Polη) (Fig. 1), a translesional polymerase that is most error-prone when it copies A and T bases [40–44]. As illustrated in Fig. 1, during normal DNA replication the resynthesis of the excised strand is carried out by high fidelity polymerases δ and ε, but at the V – and possibly S regions – Polη is recruited by mono-ubiquitylated PCNA [45,46]. The length of this patch has been estimated to be 20–30 bp [14,15,47]. It is unclear what restricts the patch to this size since in biochemical studies much longer strands of DNA can be excised by EXO1 [48,49]. Presumably nicks that are made 3′ and/or 5′ to the site of recruitment of MSH2 and MSH6 determine where EXO1 or some as yet unknown nuclease(s) initiates the excision process. While recent studies in cell free systems and in mice suggest that a latent endonuclease activity in PMS2 could generate nicks 5′ or 3′ to the mismatch [50–52], several studies have shown that the MLH1-PMS2 heterodimer (MutLα) is not required to excise the strand of DNA containing the AID-generated U:G mismatch during SHM (reviewed in [53]). This is especially peculiar since MutLα is involved in MMR as it occurs in most other circumstances *in vivo* [38,54], suggesting that it is important to prevent MLH1 and PMS2 from being recruited to the V regions. The mechanism and factors responsible for this restriction are not known.