An Evolutionary View of the Mechanism for Immune and Genome Diversity

Lucia Kato, Andre Stanlie, Nasim A. Begum, Maki Kobayashi, Masatoshi Aida and Tasuku Honjo

J Immunol 2012; 188:3559-3566; doi: 10.4049/jimmunol.1102397
http://www.jimmunol.org/content/188/8/3559

References
This article cites 65 articles, 32 of which you can access for free at:
http://www.jimmunol.org/content/188/8/3559.full#ref-list-1

Subscription
Information about subscribing to The Journal of Immunology is online at:
http://jimmunol.org/subscription

Permissions
Submit copyright permission requests at:
http://www.aai.org/About/Publications/JI/copyright.html

Email Alerts
Receive free email-alerts when new articles cite this article. Sign up at:
http://jimmunol.org/alerts
An Evolutionary View of the Mechanism for Immune and Genome Diversity

Lucia Kato, Andre Stanlie, Nasim A. Begum, Maki Kobayashi, Masatoshi Aida, and Tasuku Honjo

An ortholog of activation-induced cytidine deaminase (AID) was, evolutionarily, the first enzyme to generate acquired immune diversity by catalyzing gene conversion and probably somatic hypermutation (SHM). AID began to mediate class switch recombination (CSR) only after the evolution of frogs. Recent studies revealed that the mechanisms for generating immune and genetic diversity share several critical features. Meiotic recombination, V(D)J recombination, CSR, and SHM all require H3K4 trimethyl histone modification to specify the target DNA. Genetic instability related to dinucleotide or triplet repeats depends on DNA cleavage by topoisomerase 1, which also initiates DNA cleavage in both SHM and CSR. These similarities suggest that AID hijacked the basic mechanism for genome instability when AID evolved in jawless fish. Thus, the risk of introducing genome instability into nonimmunoglobulin loci is unavoidable but tolerable compared with the advantage conferred on the host of being protected against pathogens by the enormous Ig diversification. The Journal of Immunology, 2012, 188: 3559–3566.

Genome stability is critical for the survival of living organisms. However, complete genome stability would have prevented the evolution of organisms, as well as novel biological activities. Thus, an appropriate balance between genome stability and instability has been maintained during evolution. There are several endogenous mechanisms in addition to exogenous genome-altering mechanisms, such as viral infection, irradiation, and DNA-damaging chemicals. It is well established that high-frequency transcription enhances mutagenesis; this phenomenon is called transcription-associated mutation (TAM) (1). In addition, in the many triplet-associated diseases, such as Huntington’s disease (2), the triplets are usually located within the transcribed region of genes, and the triplet repeats are either expanded or contracted in somatic cells. Recent studies showed that topoisomerase 1 (Top1) is involved in both TAM and the triplet contraction/expansion (3–5).

Another mechanism for genome diversification in eukaryotes is mediated by meiotic recombination, which occurs in gametogenesis. Most of the mechanisms and molecules in meiotic recombination are highly conserved from yeast to humans (6). Meiotic recombination might have contributed to the scrambling of mutations and facilitate Darwinian selection.

Vertebrate immune diversity is mediated by two types of DNA-alteration mechanisms: V(D)J recombination by RAG1/RAG2 and somatic hypermutation (SHM)/class switch recombination (CSR)/gene conversion (GC) by activation-induced cytidine deaminase (AID) (7, 8). Acquired immune diversity appears to have evolved at a very early stage of vertebrate evolution, because an AID ortholog, Petromyzon marinus cytidine deaminase (PmCDA), was identified in a lamprey, P. marinus (9). In lampreys, two AID family members, PmCDA1 and PmCDA2, are expressed differentially in two types of Ag-recognizing cells that correspond to T cell and B cell precursors, which express variable lymphocyte receptors (VLR) A and VLRB, respectively. RAG1/RAG2 and V(D)J recombination appeared later in vertebrate evolution. Given their gene structure, which lacks introns, and their recognition of a specific DNA sequence, RAG1/RAG2 may have been introduced as retrotransposons, as the result of an infection (10). Thus, it is likely that the evolution of AID was the origin of acquired immune diversity.

Recent findings suggest that the SHM and CSR mediated by AID have mechanisms similar to those of meiotic recombination, TAM, and triplet contraction/expansion (3–5, 11–15). Some of these processes are shared by V(D)J recombination (8). In this review, we attempt to discuss, from an evolutionary point of view, the mechanisms for acquired immune diversity and classical genetic alterations and propose that AID adopted the basic mechanism of genome instability to generate immune diversity in lymphocytes.

General genome diversity

Genome instability. Mutations form the basis of genetic diversity and instability. Living organisms have several intrinsic mechanisms for inducing mutations in DNA, in addition to...
errors that occur during DNA replication. The best-known example is TAM, which has been observed in a wide range of organisms, including *Escherichia coli*, yeast, and mammals (1). Highly active transcription increases the frequency of various types of mutations: base replacement, deletion, and recombination. Among these, 2–5-bp deletions were recently shown to be catalyzed by Top1 (4, 5).

The 2–5-bp deletion type of TAM involves mutation hotspots containing di- or trinucleotide repeats. Efficient transcription often induces a loosening of the helical structure, resulting in the creation of excessively negative supercoiling behind the transcription machinery. Top1 is the enzyme that corrects this excessive DNA superhelix (Fig. 1A, 1B). Top1 nicks the DNA, forms a transient covalent bond with the DNA through its tyrosine residue, rotates the DNA around the helix, and then relegates the cleaved ends; this set of actions can result in an increase or decrease of the superhelix (16). Notably, the active transcription of repeat sequences is predicted to form non-B DNA structures (17, 18). When the DNA has aberrant structures like non-B DNA, rotation around the helix may be inhibited; thus, the cleavage by Top1 may be irreversible and trigger genome instability (Fig. 1C) (4, 5, 17).

Another well-known example of genome instability is associated with triplet diseases, such as Huntington’s disease, Fragile X syndrome, and myotonic dystrophy, in which triplet repeats composed of CAG, CGG, CTG, and others increase or decrease in copy number (2). These triplet repeats are usually located within a transcription unit and are probably prone to form non-B DNA structures similarly to TAM (18). Recently, triplet repeat contraction was shown to be catalyzed by Top1 (3). Taken together, these findings suggest that TAM and triplet diseases probably arise by a similar molecular mechanism, as follows: non-B DNA structures induced by the active transcription of repetitive sequences cause irreversible cleavage by Top1, triggering mutations, such as deletions and duplications (Fig. 1).

**Meiotic recombination.** Because all of the mutations that block meiotic recombination inhibit or suppress the appropriate generation of sperm or eggs, meiotic recombination is required for spermatogenesis and oogenesis (19). Meiotic recombination takes place during the first cell division. The direct outcome of meiotic recombination is the scrambling of the genetic material on the paternal and maternal chromosomes. Meiotic recombination reassembles various mutations on homologous chromosomes and facilitates the spreading of newly introduced mutations among a species population by generating new combinations of them.

The molecular mechanisms for meiotic recombination have been extensively studied using yeast genetics (15). It was recently shown that meiotic recombination requires at least two marks to determine where the DNA cleavage takes place. The DNA sequence of a loosely conserved 13mer is the cis-marking element, which is relatively abundant within the genome and usually linked to a promoter. The other mark for cleavage was recently discovered by studies on PRDM9, which recognizes the conserved 13mer DNA sequence by its zinc finger motif and carries out the trimethylation of lysine 4 of histone H3 on the adjacent chromatin (20–22). Because PRDM9 is essential for meiotic recombination, the combination of the *cis* DNA element and *trans* H3K4me3 histone modification appears to determine the initiation site of the DNA cleavage for meiotic recombination.

The enzyme that initiates DNA cleavage is well established as a type of topoisomerase II, Spo11, which forms a dimer and introduces nicking cleavage on two strands, four bases apart (23). Spo11 forms a transient covalent association with DNA, similarly to other topoisomerases. The DNA-bound Spo11 needs to be excised to generate double-stranded cleavages by the resection step involving the Mre11/Rad50/NBS1 (MRN) complex and Ctp (24). Subsequently, exonucleolytic digestion with exonuclease 1 exposes a long single-stranded region. The next event is homologous recombination, which is ini-
tiated by strand invasion of the homologous chromosome and mediated by many proteins. The Holliday structure, consisting of four strands of DNA, is stabilized by mismatch repair enzyme members, Msh4 and Msh5, which do not function as repair proteins in this situation. Finally, the Holliday junction is cut to segregate the two chromosomes, resulting in segmental exchange of the genetic material between the paternal and maternal chromosomes. Notable aspects of meiotic recombination are the marking of cleavage target chromatins by histone H3K4me3 modification and the involvement of a topoisomerase member in the DNA cleavage.

Evolution of acquired immune diversity

PmCDA1 and PmCDA2, primordial orthologs of AID, are probably the oldest enzymes responsible for Ag-specific receptor diversification in vertebrates (Fig. 2) (9, 25). PmCDA1 increases the rate of mutation and recombination when overexpressed in yeast (26). Unlike AID, the PmCDAAs are expressed constitutively, because the levels and pattern of PmCDAAs expression do not change in response to stimulation. Indeed, PmCDA1 is selectively expressed in the thymoid region where assembly of VLRA occurs, whereas PmCDA2 is expressed in VLRB-expressing lymphocytes; PmCDA1 and PmCDA2 are thought to be involved in the genetic diversification of these respective lymphocytes (27). The VLR cDNAs contain an invariant signal peptide followed by highly variable leucine-rich repeat (LRR) modules consisting of five units, which each range from 24 to 63 aa in length. The germline VLR genes only contain sequences coding for the invariant N-terminal and C-terminal portions of the VLRs separated by large noncoding sequences. However, in lymphocytes, these noncoding intervening sequences are replaced by multiple LRR-encoding modules to form a functional VLR gene. The presence of incompletely assembled VLR genes indicates that the LRR modules are put together step by step. In lamprey, 393 VLRA-related LRR cassettes and 454 VLRB-related LRR cassettes have been characterized within a continuous locus of ~2 megabase pairs surrounding their respective VLR genes (26).

The VLR gene assembly mechanism has been postulated to be of the GC type, because of the nonreciprocal insertion of LRR cassettes into the intervening sequences of the germline gene (26, 28, 29). The VLRs are highly repetitive sticky proteins, whose specificity is probably determined by their broad surface, which contains the many β sheets of LRR domains. A computational assessment of mature VLRB gene sequences predicted a potential repertoire >10^{14} AgRs (28). These striking observations indicate that the GC-type DNA rearrangement is the most primordial molecular mechanism for generating Ag-recognition receptors. Although SHM has been proposed to occur in VLRA-expressing lymphocytes, there is no evidence that VLRB diversification also occurs by SHM as the result of cytidine deaminase activity of PmCDA2 (27).

In vertebrate evolution, CSR first took place in frogs (Fig. 2). Interestingly, fish AID can complement mouse AID for CSR, although fish carry out SHM but not CSR (30–32). Considering the cytidine deaminase involvement in GC in lampreys, as well as that GC is the major mechanism of somatic diversification in chicken Ig genes (33, 34), it is likely that fish AID can carry out GC; however, this remains to be tested.

The evolutionary origin of RAGs and the recombination signal sequence is assumed to be transposon insertion (10), which must have happened in primordial jawed vertebrates, because no RAG ortholog has been identified in jawless fish, such as lampreys. RAG protein evolution must have been coupled with segregation of the Ig ancestor gene into V, D, and J segments by transposon insertion. It is striking that both H3K4me3 mark and recombination signal sequence are required for DNA cleavage by RAG1/RAG2 (8). The creation of combinational diversity by RAGs probably replaced the VLRA and VLRB system, which depends on the less efficient GC event. CDA1 has a selective disadvantage in T cells, which should avoid additional mutations after their selection in the thymus. Therefore, VLRA and CDA1 in T cells were probably lost during evolution, after RAG’s insertion into primitive vertebrates, which occurred sometime after the evolution of jawed fish (Fig. 2).

FIGURE 2. Evolution of B and T lymphocytes and their AgR diversification mechanisms. In jawless vertebrates, such as lampreys and hagfish, CDA1 and CDA2, orthologs of AID, assemble VLRA and VLRB by a GC mechanism. VLRA and VLRB are specifically expressed in B- and T-lineage cells, respectively. Sometime during the evolution of primordial jawed vertebrates, transposon insertion generated TCR and BCR genes simultaneously with site-specific recombinases RAG1/2. This enzyme mediates such an efficient and precise reassembly of AgR genes that the system using CDA1, CDA2, VLRA, and VLRB lost its advantage for generating the AgR repertoire. However, in B-lineage cells, AID that evolved from CDA2 retained its advantage because of its activity for generating Ag-induced Ab memory. It is likely that AID has maintained its GC and SHM activity in jawed vertebrates. When the Ig Cγ gene was duplicated in amphibians, AID’s GC activity was used for CSR. The DNA-cleavage mechanism in CSR is very similar to that in SHM and GC. The recombination-specific activity in CSR is probably also similar to that in GC, because both depend on the pairing of DNA segments in adjacent loci. The genes, enzymes, and genetic events involved in the primordial immune diversity and their descendants are in green. Myrs, million years; Rb, recombination.

The Journal of Immunology 3561
In contrast, in B-lineage cells, even after the evolution of the RAGs and Ig genes, AID maintained its selective advantage by being required for Ag-induced affinity maturation and Ab memory through its GC mechanism, which probably coexisted with SHM. Thus, AID carries out GC and SHM in the Ig locus of chicken and other vertebrates. Mouse and human AID can exhibit GC activity in chicken DT40 B cells (35, 36), although it is not clear whether mouse or human AID has GC capability in vivo.

**Similarities between mechanisms for immune and genome diversity**

AID has two functions: 1) the DNA cleavage that is required for CSR, SHM, and GC and 2) the activity specific to CSR and probably GC. The N-terminal portion of AID is required for DNA cleavage, whereas the C-terminal portion is required for CSR, as described later (35, 37–39). Two contrasting hypotheses were put forward to explain how AID induces DNA cleavage in the target DNA; the initially proposed RNA-editing hypothesis assumed that AID edits RNA to generate a DNA cleaving enzyme, whereas the much straightforward DNA-deamination hypothesis postulated that AID directly deaminates target DNA to initiate DNA cleavage. Recent results from our studies revealed striking similarities between mechanisms for immune diversity and genome diversity.

**Involvement of Top1 in DNA cleavage.** Several lines of evidence indicate that Top1 is involved in the DNA cleavage in CSR and SHM. First, AID activation reduces Top1 protein level; interestingly, this reduction is not seen in a loss-of-function mutant of AID (12). Second, artificial reduction of Top1 by small interfering RNA augments DNA-cleavage activity, SHM, and CSR by AID (11, 12). Furthermore, in vivo, the haplo-insufficiency of Top1 (by the cross of Top1<sup>fl/fl</sup> with CAG-cre mice) dramatically augments SHM in B cells isolated from Peyer’s patches (gut follicular structures in which germinal centers are continuously induced by bacterial stimulation) (11). Thus, it is clear that a reduction in Top1 correlates with enhanced DNA cleavage and a subsequent increase in SHM or CSR.

How can AID reduce the Top1 protein level? We proposed that the N-terminal portion of AID may interact with a putative cofactor N that recognizes a precursor of the microRNA for editing (12). The edited microRNA is assumed to bind together with the RNA-induced silencing complex to the 3’ untranslated region of Top1 mRNA (Fig. 3A). This binding would suppress the translation of Top1 mRNA, causing reduction of the protein. In support of this hypothesis, specific depletion of Dicer from AID-expressing B cells drastically reduced SHM and CSR (40). Importantly, the observed reduction in switching, SHM, and affinity maturation in Dicer-deficient B cells was not due to cell death, because it could be still observed in B cells that were rescued from apoptosis by the deletion of Bim, a proapoptotic member of the Bcl-2 family (40).

**Non-B DNA structures as a mark for cleavage target.** As mentioned above, Top1 is involved in correcting the excessive positive and negative supercoils that accumulate before and after the transcription machinery, respectively. However, when Top1 is reduced, it cannot appropriately correct the excessive supercoil, and, as a result, more non-B DNA forms. An important aspect of this mechanism is that the appropriate Top1 level is determined by a balance with the efficiency of transcription. The augmented transcription in TAM causes a relative reduction in Top1 at the locus.

When Top1 tries to correct the excessive supercoil by nicking, the non-B form DNA prevents rotation of the DNA around the helix, and Top1 is stuck, bound to the DNA (Fig. 1). To generate cleaved DNA ends, the covalently associated Top1 is probably removed by several resection enzymes, such as the MRN complex and Ctp, which are used in meiotic recombination. Consistent with this scenario, when resection of the Top1–DNA complex is blocked by the Top1 intercalating inhibitor Camptothecin (30 nM), the double-stranded break formation, SHM, and CSR are inhibited (11, 12). It is well known that both the S and V regions contain direct or inverted repeats, which are prone to form non-B structures (17, 41, 42).

Previously, a large number of AID target loci was proposed by the experiments with chromatin immunoprecipitation and microarray or genome-wide sequencing (43). These experiments were based on the assumption that AID-interacting molecules are located at the AID-cleavage sites. Unfortunately, however, there were no consensus features among a large number of proposed AID targets. To search for common properties of highly efficient targets of AID, we carried out a new approach to identify the cleaved loci. We used human Burkitt lymphoma BL2 cells expressing the C-terminally truncated AID fused with the estrogen receptor hormone-binding domain, which were briefly activated (3 h) with tamoxifen (44). AID-induced DNA cleaved ends were labeled by a biotin oligonucleotide and enriched by streptavidin beads. A library of DNA fragments containing the cleaved ends was amplified and analyzed by whole-genome sequencing and microarray analyses. The candidates were further confirmed by quantitative PCR quantitation of their actual cleavage sites in the original cleaved end-enriched library. Finally, we confirmed that mutations accumulated in the new loci as frequently as did Ig loci. Thus, we identified four hypermutating loci: MALAT1, SNHG3, BCLI7A, and CUX1 (44). All of these loci are flanked by abundant repetitive sequences, which are prone to form non-B DNA structures.

**Histone modification H3K4me3 as a mark for DNA cleavage.** In addition to the non-B DNA structure, chromatin modification, especially H3K4me3, was shown to be essential for DNA cleavage of the AID target. Knockdown of the histone chaperones facilitates chromatin transcription complex or Sp56 or the elongation factor Sp5 blocked the generation of H3K4me3 and DNA cleavage at the AID target regions without inhibition of S region transcription, and yet it blocked AID-induced cleavage (13) (A. Stanlie, N.A. Begum, H. Akiyama, and T. Honjo, submitted for publication; N.A. Begum, A. Stanlie, H. Akiyama, and T. Honjo, submitted for publication). Knockdown of PTIP abolished CSR with simultaneous blockade of H3K4me3 and S region transcription, leaving the possibility that CSR inhibition could be due to loss of target transcription (45). H3K4me3 accumulation was shown in both the V and S regions of the Ig locus. The nucleosomes of the newly identified AID targets (MALAT1 and others) also have the H3K4me3 mark (44). Another well-known target of AID, the Myc locus, is...
flanked by repetitive sequences and marked by H3K4me3 (46, 47). Therefore, we believe that cis marking by non-B structural formation and trans marking by H3K4me3 on chromatin serve as marks for recruiting Top1 to the cleavage target, probably with the help of other proteins (Fig. 4). Remarkably, H3K4me3 is involved in marking cleavage targets in meiotic recombination, as well as V(D)J recombination.

Evolutionary consideration of the DNA-deamination model

The DNA-deamination model assumes that AID directly binds DNA and deaminates cytosine. The conversion of C to U generates a mismatched U/G base pair, followed by DNA cleavage introduced by the base excision repair or mismatch repair pathway (48). There is much evidence for the DNA-deamination model, the most straightforward of which is the in vitro deamination of DNA by recombinant AID (49–51).

FIGURE 3. The RNA-editing hypothesis for AID. (A) AID’s N-terminal domain is required for DNA cleavage. This activity probably depends on a putative cofactor N that interacts with the N-terminal region of AID and captures a precursor of microRNA. AID edits the microRNA precursor in the nucleus, and the edited microRNA is processed and interacts with the Top1 mRNA 3’ untranslated region in a complex with the RISC protein. This interaction inhibits the translation of Top1 RNA, resulting in a reduction in the Top1 protein. The reduction in Top1 protein enhances the amount of non-B form DNA in the S and V regions, which are actively transcribed. The formation of the unusual structure in the S and V regions induces irreversible cleavage by Top1. (B) The C-terminal domain of AID is involved in editing the mRNA encoding a putative bending factor. The mRNA is captured by a putative cofactor C, which interacts with the C-terminal region of AID. The edited mRNA is translated into a new DNA-bending factor. This protein (denoted by a yellow ring) is required to pair two switch regions or two V regions for CSR and GC, respectively, by DNA bending. The stabilization of the DNA pairing probably requires additional proteins, such as UNG and Msh2/6.

FIGURE 4. Common features between acquired immune diversity and genetic instability and diversification. (A) Meiotic recombination, V(D)J recombination, and CSR/SHM all depend on two marks on the targets: H3K4me3 on chromatin (trans) and unique DNA structure (cis). The cleaving enzymes are different. However, both meiotic recombination and CSR use members of the topoisomerase family. The non-B structure formation in CSR is enhanced by the AID-dependent decrease in Top1. (B) Genome instability mechanisms, such as those in TAM and triplet diseases, depend on the strong activation of transcription, which induces a relative Top1 deficiency. Non-B DNA structures might also be enhanced by the relative Top1 deficiency. Top1 is also a DNA-cleaving enzyme that probably recognizes non-B structures. Although it has not been assessed whether this target is also marked by H3K4me3, genome instability shares similar mechanisms with AID-induced acquired immune diversity. Rb, recombination; Set1/MLLs, histone methyltransferases.
In this system, ssDNA is a much better substrate than dsDNA for AID (52). In addition, nucleosomal DNA is a less efficient target for DNA deamination (53). However, the overexpression of AID in E. coli or yeast introduces preferential C- to-T or G-to-A mutations. Another piece of evidence is the requirement for uracil-DNA glycosylase (UNG), which removes U from the U/G mismatch, for CSR and GC, although UNG inhibits SHM (54, 55). All of these mechanisms are unique and are not known to be involved in any types of genetic alterations. On the contrary, both base excision repair and mismatch repair enzymes are required for the prevention of genetic alterations rather than genome instability.

Although broadly accepted, the DNA-deamination model has shortcomings that have been extensively discussed in other reviews (7, 56). Therefore, we will list only a few points here. First, the in vitro and in E. coli DNA-deamination activities are catalyzed by a bona fide RNA-editing enzyme APOBEC 1, suggesting that these artificial assays do not provide evidence for in vivo DNA deamination by AID (57, 58). In fact, APOBEC 1 requires a cofactor ACF for its RNA-editing activity in vivo, as well as in vitro (59). To explain the AID target specificity, the DNA-deamination hypothesis postulates the presence of guiding factors that physically interact with AID. A large number of AID-interacting proteins has been found: PKA, PTEN, RNA exosome complex, RPA, Trim28, MDM2, Pol II, Sp16, Sp5, DNA-PK, and 14-3-3 (43). However, it is difficult to explain how such a large number of proteins can specify DNA targets, because most of them are rather ubiquitous proteins.

Evolutionary consideration of the CSR-specific role of the C-terminal region of AID

The C-terminal region of AID is highly conserved among vertebrates, suggesting that it has an important function (31, 60). This region of AID has two activities: the nuclear export signal and the CSR-specific activity. This is because C-terminal mutants retain the in vitro DNA-deamination activity and in vivo SHM activities in both the V and S regions but lose CSR activity (35, 38, 39). Thus, it is likely that the CSR-specific function of the C-terminal domain of AID is independent from the DNA-deamination activity.

Several hypotheses have been proposed for the function of AID’s C-terminal domain, including stabilization of the AID protein or protection from cell death induced by double-strand breakage (61–63). However, these hypotheses cannot explain that a C-terminal mutation causes a loss of CSR but less of a defect in SHM in hyper-IgM syndrome type 2 patients (39). In addition, C-terminally truncated AID mutants fused with the estrogen receptor hormone-binding domain, which stabilizes the protein, show a similar loss of CSR with intact or hyper-SHM, making the stability hypothesis less likely (37).

The C-terminal domain of AID is responsible for binding poly(A) RNA through a putative cofactor protein (64). Consistent with this idea, CSR depends on de novo protein synthesis (65). It was proposed that synopsis formation may be required for appropriate pairing of cleaved ends of two separate S regions (66–68). These observations led us to speculate that AID edits an mRNA captured by cofactor C, as well as that translation of the edited mRNA generates a novel protein required for the synopsis formation between cleaved ends on the same chromosome (Fig. 3B). This protein, which we call “bending factor,” may act with UNG and Msh2/6, which are also required for GC (54, 55). Because both GC and CSR require recombination between cis loci, UNG and Msh2/6 might be involved in pairing the cis loci required for both GC and CSR. Because the evolution of AID originated from a role in GC, AID must have had two activities (DNA cleavage and cis pairing) from the beginning. If so, it is not surprising that fish AID can exhibit CSR activity in mouse cells. To test this hypothesis, it is important to re-examine whether the C-terminal domain of AID is also required for GC. Although a C-terminal truncation mutant of AID was reported to show the same level of GC and SHM as the wild-type AID (35), the data appear to contradict the observation of Barreto et al. (35) that this AID mutant reverted a stop codon mutation in the VA-chain gene of DT40 cells almost 10 times more frequently than did the wild-type AID.

Conclusions

Because nothing in biology makes sense except in the light of evolution (69), we compared the mechanisms for various types of genome diversification. Surprisingly, recent research on AID function revealed that immune and genome diversification mechanisms share many features with respect to cleaving enzymes and target markers. Top1 is used in AID-induced DNA alterations (CSR and SHM), as well as in TAM deletion and triplet contraction/expansion (Fig. 4). Meiotic recombination uses a type of topoisomerase, Spo11, for DNA cleavage. The cis markings for CSR, SHM, TAM, and triplet contraction/expansion all appear to depend on a non-B DNA structure. H3K4me3 on chromatin of the cleavage target is used as a trans mark for CSR, SHM, meiotic recombination, and V(DJ) recombination. It is worth examining whether there is some posttranslational modification of histone in targets of TAM and triplet contraction/expansion, because transcription is mandatory for both events. We propose that AID evolution in the primordial vertebrates took advantage of the pre-existing genome instability mechanism for DNA cleavage and recombination for accomplishing receptor diversity. Nonetheless, the risk of introducing genome instability into non-Ig loci is tolerable compared with the advantage of the protection against pathogens by the enormous immune diversification by AID.

Acknowledgments

We thank Drs. M. Yamamoto, S. Takeda, H. Nagaoka, K. Kinoshita, M. Muramatsu, and S. Fagarasan for critical reading and comments on the manuscript, as well as Y. Shiraki and K. Fukui for manuscript preparation.

Disclosures

The authors have no financial conflicts of interest.

References


