

Molecular Mechanisms of Antibody Somatic Hypermutation

Javier M. Di Noia¹ and Michael S. Neuberger²

¹Institut de Recherches Cliniques de Montréal, H2W 1R7 Montréal, Québec, Canada; email: javier.di.noia@ircm.qc.ca

²MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, United Kingdom; email: msn@mrc-lmb.cam.ac.uk

Annu. Rev. Biochem. 2007. 76:1–22

First published online as a Review in Advance on February 28, 2007

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

This article's doi:
10.1146/annurev.biochem.76.061705.090740

Copyright © 2007 by Annual Reviews.
All rights reserved

0066-4154/07/0707-0001\$20.00

Key Words

activation-induced deaminase, class switch recombination, immunoglobulin gene diversification, uracil excision

Abstract

Functional antibody genes are assembled by V-D-J joining and then diversified by somatic hypermutation. This hypermutation results from stepwise incorporation of single nucleotide substitutions into the V gene, underpinning much of antibody diversity and affinity maturation. Hypermutation is triggered by activation-induced deaminase (AID), an enzyme which catalyzes targeted deamination of deoxycytidine residues in DNA. The pathways used for processing the AID-generated U:G lesions determine the variety of base substitutions observed during somatic hypermutation. Thus, DNA replication across the uracil yields transition mutations at C:G pairs, whereas uracil excision by UNG uracil-DNA glycosylase creates abasic sites that can also yield transversions. Recognition of the U:G mismatch by MSH2/MSH6 triggers a mutagenic patch repair in which polymerase eta plays a major role and leads to mutations at A:T pairs. AID-triggered DNA deamination also underpins immunoglobulin variable (IgV) gene conversion, isotype class switching, and some oncogenic translocations in B cell tumors.

Contents

INTRODUCTION.....	2	Partners of AID	11
HISTORICAL BACKGROUND ...	3	TARGETING OF AID AND THE	
THE ROLE OF SOMATIC		LINKAGE TO	
HYPERMUTATION IN		TRANSCRIPTION	11
ANTIBODY		PROCESSING THE	
DIVERSIFICATION	3	AID-GENERATED U:G	
INTRINSIC FEATURES OF		LESION.....	12
SOMATIC		Uracil Excision by UNG.....	12
HYPERMUTATION	5	Recognition of the U:G Mismatch	
TWO PHASES OF SOMATIC		by MSH2/MSH6.....	13
HYPERMUTATION	6	UNG and MSH2/MSH6 Provide	
THE FIRST PHASE: DNA		Alternative Pathways for	
DEAMINATION.....	6	Processing the U:G Lesion	13
DNA DEAMINATION ALSO		Replication Over the	
TRIGGERS GENE		UNG-Generated Abasic Site ..	14
CONVERSION AND CLASS		INTRODUCTION OF	
SWITCH RECOMBINATION ..	7	MUTATIONS AT A:T PAIRS....	14
EVIDENCE FOR THE DNA		Relationship to Conventional	
DEAMINATION SCHEME	7	Mismatch Repair	14
CHARACTERISTICS OF AID	9	Mechanism of Mutagenesis at	
AID as a Polynucleotide		A:T Pairs	15
Deaminase.....	9	MUTATION VERSUS REPAIR	15
Subcellular Localization of AID ...	10	CONCLUSION	16

INTRODUCTION

Vertebrates are able to produce a vast repertoire of antibody molecules in order to combat infection. The number of different antibodies that a human can produce during his lifetime (estimated to be in excess of 10^9) greatly exceeds the coding capacity of the inherited genome. Instead, the size of the expressed antibody repertoire owes much to somatic gene diversification processes.

The DNA encoding the antigen-combining portion of the antibody is assembled by gene rearrangement during the early phase of B lymphocyte development [reviewed in (1)]. Integration of the immunoglobulin (Ig) variable (V), diversity (D) and joining (J) segments allows the production of a primary repertoire of antibody specificities. The diversity of this repertoire derives both from the

choice of germ line V, D, and J gene segments available for integration (combinatorial diversity) as well as from imprecision at the sites of V-D-J gene integration (where deletion or nontemplated insertion of nucleotides generates junctional diversity).

The size of the primary repertoire varies from species to species and locus to locus, depending on the number of germ line V, D, and J segments as well, for example, as on the availability of terminal deoxynucleotidyl transferase (which catalyzes nontemplated nucleotide insertion). Thus, whereas chickens have only one or a few functional V, D and J segments in their Ig loci and generate scarcely any diversity through gene rearrangement (2), the rearrangement process in man and mouse is thought to provide around 10^5 – 10^6 different antibody specificities.

In no species, however, is the repertoire of antibodies generated through gene rearrangement large enough to provide high-affinity antibodies (e.g., K_d 10^9 M^{-1}) against the wide range of antigens that an animal might encounter. The generation of antibody diversity therefore depends in a major way on processes that occur following V-D-J integration. These processes are either somatic hypermutation (in which multiple, single nucleotide substitutions are introduced in and around the productively rearranged V(D)J gene segment) or gene conversion (in which nonfunctional V pseudogenes located close to the productively rearranged V(D)J segment are used as templates for its patchwork diversification). Here we focus on somatic hypermutation (SHM), which accounts for diversification of the primary antibody repertoire in man and mouse, although, as we discuss below, gene conversion (which plays a central role in postrearrangement diversification in chickens) is triggered by a very similar process [reviewed in (3)].

HISTORICAL BACKGROUND

The clonal selection theory, which was put forward 50 years ago, underpins much of our thinking about the mechanism by which specific antibody responses are elicited (4). Burnet noted that “The theory requires at some stage in early embryonic development a genetic process for which there is no available precedent. In some way we have to picture a ‘randomisation’ of the coding responsible for part of the specification of gamma globulin molecules . . .” Two years later, Lederberg (5) proposed that this randomization of the coding sequence might be achieved by “the precursors of antibody-producing B cells undergoing a high rate of spontaneous mutation during their lifelong proliferation.”

However, during the ensuing years it became evident that antibody diversity was largely restricted to the amino terminal portion of Ig polypeptide chains (6–8) making it unlikely that the diversity was achieved by

the lymphocyte undergoing a period of hypermutation across its entire genome. In light of this, Brenner & Milstein (9) resurrected Lederberg’s 1959 proposal of SHM but suggested that targeted antibody diversity might be achieved by some form of localized DNA synthesis triggered by error-prone repair of a lesion caused by some unidentified DNA-cleaving enzyme.

Several other mechanisms that might account for antibody diversity were also put forward around this time with many of these being discussed at the 1967 Cold Spring Harbor meeting on Antibodies (10). With the advent of molecular cloning in the mid-1970s, it became clear that functional antibodies were assembled by somatic rearrangement of germ line Ig gene segments and that this could account for much of antibody diversity. However, even back in 1970, a comparison of the amino acid sequences of mouse Ig λ light chains led Weigert et al. (11) to conclude that their derivation was most likely explained by spontaneous mutation of a V λ germ line progenitor. This interpretation was buttressed by the subsequent characterization of the genomic mouse Ig λ 1 locus (12).

The full extent of the contribution of SHM depended on the ability to compare the sequences of expressed Igs with those available in the germ line. This was facilitated either by the use of tagged transgenes (13) or by increasing knowledge of organismic genome sequences. Such studies revealed that SHM makes a major contribution to antibody diversity and, indeed, a dominant contribution to antibody affinity maturation and the diversity of the secondary antibody repertoire in man and mouse [reviewed in (14, 15)].

THE ROLE OF SOMATIC HYPERMUTATION IN ANTIBODY DIVERSIFICATION

SHM in man and mouse takes place following V-D-J recombination and occurs both in a different part of the body as well as in a different phase of the immune response. Thus,

SHM: somatic hypermutation

CSR: class switch recombination

AID: activation-induced deaminase

gene rearrangement occurs during B cell development in omentum, fetal liver, or adult bone marrow and is used to produce a primary repertoire of antibodies, which is generated prior to encounter with foreign antigens. This provides a first line of defense by way of IgM antibodies that usually exhibit only low binding site affinity for the antigen, although binding of such IgMs to their targets can be sufficient to trigger physiological responses.

Binding of the antigen to B cells that express cognate (through low affinity) IgM antibodies causes these B cells to undergo rapid proliferation forming structures within secondary lymphoid organs that are termed germinal centers. It is within these germinal center B cells that SHM occurs, and an iterative alternation of SHM and antigen-mediated selection lead to antibody affinity maturation [reviewed in (15–17)]. Furthermore, in addition to the introduction of point mutations in the rearranged immunoglobulin variable (IgV) segment, regions of the Ig heavy-chain locus can be targeted for class switch recombination (CSR), a process of recombinational deletion that removes the exons of the IgM constant region, bringing the functional VDJ segment into proximity with the exons of downstream Ig constant (IgC) regions (**Figure 1a**). This recombination among the IgC regions (which, as discussed below, is triggered by a similar process to SHM and gene

conversion at the IgV) allows the switch from IgM to IgG, IgA, or IgE [reviewed in (18)].

Thus, over a period of a few days following first encounter with antigen, the genes that encode low-affinity IgM antibodies specific for the antigen are transformed into genes encoding high-affinity antibodies, typically of an IgG, IgA, or IgE isotype. This neo-Darwinian, real-time evolution of high-affinity binding sites underpins the production of specific antibodies, which gives protection against subsequent encounter with the same antigen (immunological memory) and also provides a biological paradigm for the rapid evolution of protein-binding sites.

Although SHM therefore fulfills a central role in antibody affinity maturation in man and mouse, it can also function in the generation of the primary (as opposed to solely secondary) antibody repertoire in, for example, sheep and possibly lower vertebrates (19, 20). The key enzymes for triggering V-D-J gene integration (the RAG1/RAG2 recombinase) and for triggering IgV SHM activation-induced deaminase (AID) both appear to have arisen around the beginning of vertebrate evolution (21, 22), but it is presently not possible to say whether SHM arose before or after gene rearrangement as a mechanism for diversifying the repertoire of antigen-neutralizing proteins in the body fluids of living organisms.

Figure 1

Antibody gene diversification. (a) Schematic representation of the Ig gene diversification processes as exemplified in the immunoglobulin heavy chain (IgH) locus. The drawing is not to scale. The double slash indicates that the C μ and C γ regions are not contiguous. Curved lines with arrows indicate gene rearrangement events; straight lines with arrows indicate the positions of transcription promoters. The regions where somatic hypermutation accumulates are indicated by curves reflecting the frequency of mutation (see Reference 94). The VDJ region was expanded as a histogram to illustrate the existence of intrinsic mutational hot spots and of the nonrandom nature of mutational targeting. CDRs (complementarity-determining regions) constitute the parts of the antibody IgV region that are most implicated in antigen contact. (b) Table showing the typical percentages of the mutations that are due to each possible type of base substitution during somatic hypermutation in mice and humans. Transition mutations are highlighted in red, and the total percentage of mutations that are targeted to each of the four bases is indicated on the right. Abbreviations: A, deoxyadenosine; C, deoxycytidine; D, diversity segment; G, deoxyguanosine; J, joining segment; Pur, purines; Pyr, pyrimidines; T, deoxythymidine; V, variable segment.

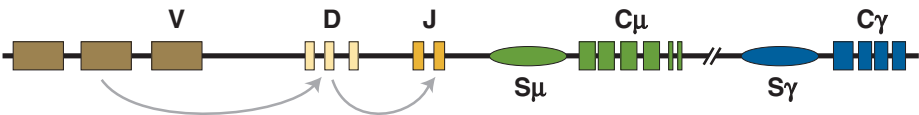
INTRINSIC FEATURES OF SOMATIC HYPERMUTATION

SHM of IgV genes occurs during only a narrow window of B cell development, introducing single nucleotide substitutions in a stepwise manner at a frequency of around 10^{-3} per base pair per generation into a do-

main that starts some 150 nucleotides downstream of the IgV promoter and extends over about 2 kb (15–17) (see **Figure 1**). All four bases can be targeted for mutation, and in man and mouse, C:G pairs and A:T pairs are targeted with approximately equal frequency (**Figure1b**).

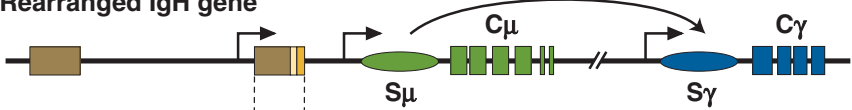
a

(i) Germ line IgH gene



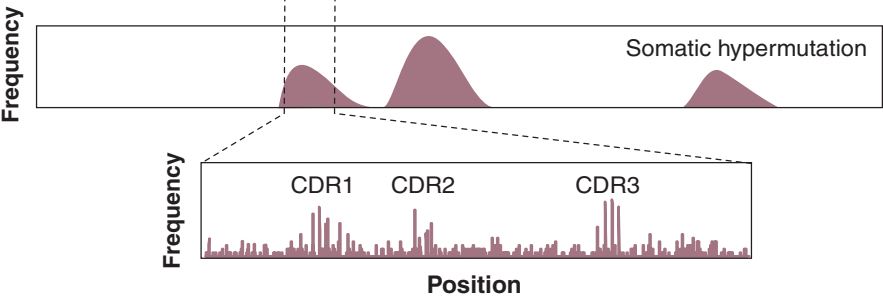
VDJ recombination

(ii) Rearranged IgH gene



Class switch recombination

(iii) Distribution of mutations



b

From \ To	Pyr		Pur		
	T	C	G	A	
Pyr	T		4	2	13
	C	16	2	4	22
Pur	G	7	6	15	28
	A	3	12	22	37

The targeting of individual bases is, however, not random. Mutations at A:T pairs are more likely to occur if the A (rather than the T) is located on the coding strand. This results in twofold more mutations occurring at A (on the coding strand) compared to T. No such strand discrimination is apparent with regard to mutations at C:G pairs (23, 24) (see **Figure 1**). Mutations are also nonrandom in that preferred mutational hot spots and cold spots can be discerned within the target region (25–27). Many of the major mutational hot spots at C:G pairs occur within a WRCY consensus (where W = A/T, R = A/G, and Y = C/T) with AGCT being a preferred embodiment. Although most major hot spots conform to the consensus, not all consensus sequences within the IgV mutation domain are sites of mutational hot spots.

Interestingly, the sequences of germ line IgV genes exhibit a biased codon usage such that certain regions (especially those implicated in antigen binding or maturation of antigen affinity) are intrinsically more mutable than others (e.g., structurally important framework residues) (28–30). Thus, germ line IgV sequences may have evolved so as to optimize the targeting of amino acid replacements during SHM in order to achieve efficient antibody maturation.

The nature of the substitutions is also nonrandom with transitions (exchange of a purine for a purine or of a pyrimidine for a pyrimidine) accounting for roughly half the nucleotide substitutions rather than the one third that would be expected on a random basis.

TWO PHASES OF SOMATIC HYPERMUTATION

Although SHM in man and mouse targets C:G and A:T pairs with roughly equal favor, it was discovered that mice deficient in the mismatch recognition protein MSH2 exhibit an altered spectrum of mutations with a marked reduction specifically in mutations at A:T (31–33). This led to the proposition (33)

that somatic mutation occurs in two phases: a first phase that targets C:G pairs and a second phase (dependent on MSH2-mediated recognition of a lesion generated during the first phase) that introduces substitutions at A:T pairs. In fact, it is this first C:G-biased phase of somatic mutation that is likely to be the predominant phase of mutation occurring in hypermutating B cell lines and possibly also in vivo in the frog (34–37).

THE FIRST PHASE: DNA DEAMINATION

Evidence accumulated over the past few years has revealed that SHM and gene conversion at the IgV as well as switch recombination of the IgC are all triggered by targeted deamination of deoxycytidine residues [reviewed in (38)]. This deamination (which changes deoxycytidine into deoxyuridine and consequently transforms C:G pairs into U:G mispairs) is catalyzed by the enzyme AID, an enzyme which is specifically expressed in activated B lymphocytes (39). As discussed below, the precise spectrum of nucleotide substitutions introduced during somatic mutation (as well as the induction of IgV gene conversion and IgC switch recombination) largely depends on the way in which the initiating U:G lesion is recognized, processed, and resolved.

AID was identified in a subtractive hybridization screen for cDNAs induced during cytokine-mediated activation of switch recombination in a mouse B cell line (39). It was then discovered that both SHM and switch recombination were ablated in vivo in AID-deficient mice and humans, which are viable and show no other obvious phenotypes except for a hyper-IgM syndrome (40, 41). The demonstrations that AID was essential for IgV gene conversion in the context of an immortalized chicken B cell line (42, 43) and that introduction of AID could drive IgV somatic mutation or IgC switch recombination in mouse cell lines (44, 45) indicated that AID was closely involved in the antibody

diversification process rather than simply being implicated in bringing the B cell to a suitable stage of differentiation.

Insight into the molecular function of AID was provided because it exhibits striking homology to APOBEC1. Thus, APOBEC1 is the catalytic component of a tissue-specific RNA editing complex, which deaminates C₆₆₆₆ to U in the RNA that encodes the apolipoprotein B polypeptide. This leads to the creation of a premature stop codon and the synthesis of the short form of apolipoprotein B polypeptide [reviewed in (46)]. Although it was initially suggested that AID would also, by analogy, somehow function in antibody diversification through editing of some unidentified RNA molecule (40), it was later proposed that AID might rather act by directly deaminating C residues within Ig locus DNA (47).

DNA DEAMINATION ALSO TRIGGERS GENE CONVERSION AND CLASS SWITCH RECOMBINATION

DNA deamination was an attractive hypothesis to explain antibody diversification in view of the evidence already indicating that SHM/CSR (40, 41, 48, 49) and somatic hypermutation/gene conversion (50–52) might share similar initiating events with, as discussed above, the data pointing to the initiating event in somatic mutation being the introduction of a lesion specifically at C:G pairs (33).

Although somatic hypermutation, gene conversion, and switch recombination are distinct processes, the DNA deamination scheme (47) envisages that they are all triggered by AID-mediated DNA deamination within the Ig loci. In this scheme (see **Figure 2**), DNA synthesis over the AID-generated U:G lesion or over the abasic site that is formed by excision of the uracil through the action of a uracil-DNA glycosylase is seen as leading to transitions and transversions at C:G pairs. Mutations at A:T pairs are envisaged as re-

sulting from a mutagenic patch repair that is triggered by MSH2/MSH6-mediated recognition of the U:G lesion.

In the case of gene conversion as occurs in the chicken IgV locus, it is envisaged that the AID-generated U:G lesion in the rearranged IgV(D)J is resolved by a recombinational repair process using the neighboring IgV pseudogenes as templates. The choice of a gene conversion as opposed to SHM outcome is probably affected by the availability of the adjacent IgV pseudogenes, the balance of available DNA repair factors, and, possibly, the stage of the cell cycle in which the lesion is resolved. Modifying any of these factors might therefore affect the final outcome. Thus, chicken B cells, which are either deficient in one of the RAD51 paralogues (e.g., XRCC2 or XRCC3, implicated in recombinational repair) or which carry a deletion spanning the IgV pseudogenes, largely diversify their immunoglobulin genes by a hypermutation (comprising base substitutions at C:G pairs) at the expense of gene conversion (52, 53).

With regard to switch recombination, AID-mediated deamination in the vicinity of both donor and acceptor switch regions is envisaged as leading to DNA double-strand breaks that promote the switch-associated deletion of the intervening region. Again, deficiencies in different DNA repair proteins affect the efficiency and features of the switch junctions differently, probably reflecting that DNA double-strand breaks at the S regions can be dealt with in many ways [reviewed in (18)].

EVIDENCE FOR THE DNA DEAMINATION SCHEME

This DNA deamination scheme made two major types of prediction. One was that, unlike any previously identified enzyme, AID should have the ability to specifically catalyze C to U deamination in DNA. The other was that a U:G lesion in the Ig locus DNA constitutes a central intermediate in antibody gene

Abasic site: a position in DNA where the nucleotide base has been removed, leaving the deoxyribose-phosphate backbone intact

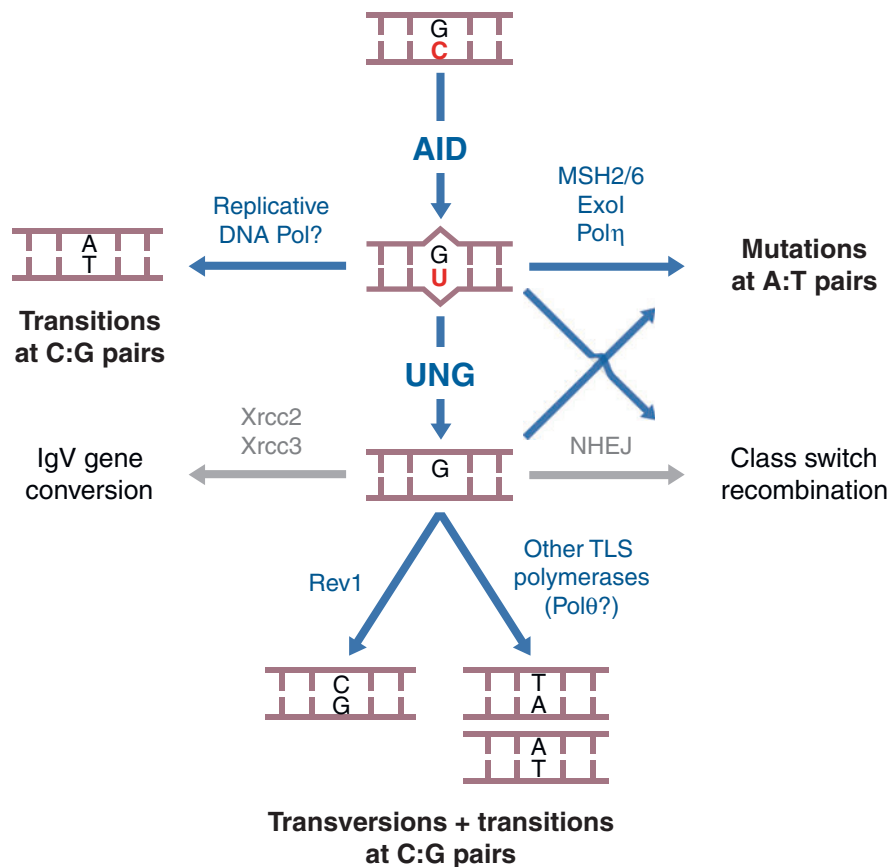


Figure 2

DNA deamination model of immunoglobulin gene diversification, emphasizing somatic hypermutation and indicating some of the key enzymes implicated in each pathway. Abbreviations: AID, activation-induced deaminase; ExoI, mismatch repair exonuclease I; IgV, immunoglobulin variable region; NHEJ, nonhomologous end joining; MSH2/6, mismatch recognition proteins; Pol, polymerase; Rev1, Y-family DNA polymerase involved in DNA damage tolerance; TLS, translesion synthesis; UNG, uracil-DNA glycosylase; Xrcc2 and Xrcc3, Rad51 paralog proteins involved in homologous recombination repair.

diversification pathways: Modifying or ablating enzymes involved in processing U:G lesions should therefore significantly perturb antibody gene diversification pathways. Both types of prediction have been confirmed in a variety of studies.

Thus, both bacterial genetic studies as well as biochemical studies of recombinant AID have confirmed that it is able to deaminate cytosine in DNA (47, 54–59). Indeed, AID shows a preference for C residues that lie within a WRC consensus, suggesting that

the WRCY hot spots observed during IgV somatic mutation in vivo could result from AID oligomers acting preferentially on targets comprising overlapping WRC motifs on opposite DNA strands (57, 60).

With regard to genetic evidence that the U:G lesion is an intermediate in antibody diversification, a deficiency in the major uracil-excision enzyme (UNG) perturbs somatic mutation, gene conversion, and switch recombination exactly as predicted by the scheme (61–66). Similarly, the second phase

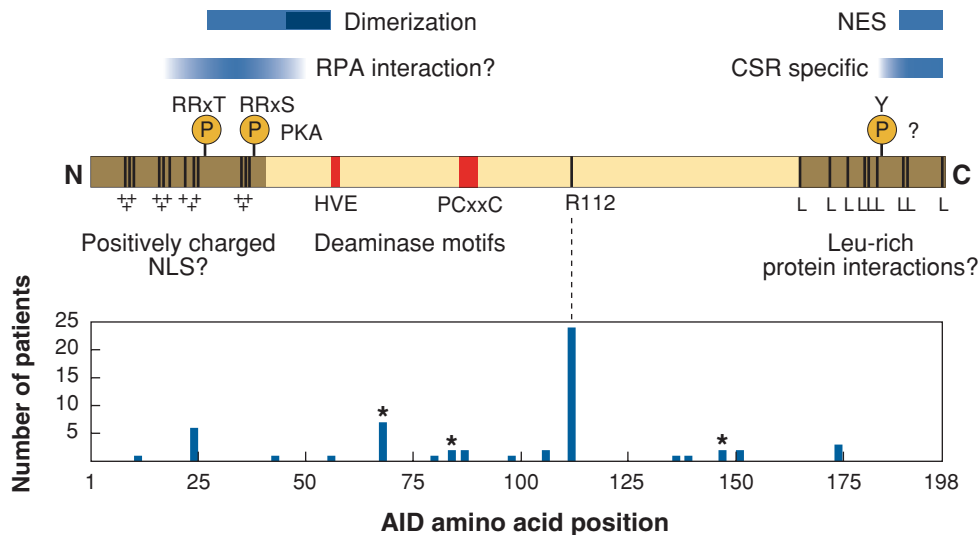


Figure 3

Organization of activation-induced deaminase (AID) functional domains. AID is drawn to scale with the N- and C-terminal ends indicated by N and C, respectively. The positions of positively charged amino acids in the N terminus and leucine (Leu) residues in the C terminus are indicated by black vertical lines with + and L symbols, respectively. Both regions are in brown with the suggested function indicated (NLS, nuclear localization signal). Red boxes indicate the deaminase motifs that are likely implicated in Zn coordination, with the conserved amino acids typed below (H, histidine; V, valine; E, glutamic acid; P, proline; C, cysteine). In the upper part, blue bars indicate the approximate position of known AID functional domains. Abbreviations: RPA, replication protein A; NES, nuclear export signal; CSR, class-switch recombination. Gold circles indicate consensus phosphorylation sites for protein kinase A (PKA) as well as a tyrosine (Y) that has been found to be phosphorylated in vivo (R, arginine; T, threonine; S, serine) (88, 89). In the lowest part, a histogram shows that the mutation of AID arginine 112 (to histidine or cysteine) is particularly common in hyper-IgM patients, but the mechanisms by which this substitution disrupts AID function is not known. The compilation of hyper-IgM AID mutations is taken from <http://bioinf.uta.fi/AICDABase/>. Patients reported with hyper-IgM syndromes owing to an AID deficiency in which a given position bears a stop codon are indicated by an asterisk. Insertion and deletion mutations were omitted in this figure.

of somatic mutation (generating mutations at A:T pairs) is disrupted through a deficiency in MSH2 (33) or MSH6 (67), consistent with it being the MSH2/MSH6 heterodimer that recognizes U:G mismatches (68).

CHARACTERISTICS OF AID

Progress in elucidating the three-dimensional structure of AID has been hindered by the difficulty in producing enough active, recombinant protein. Nevertheless, sequence comparisons with other deaminases and mutagenesis studies have provided some infor-

mation regarding functional domains in AID (see **Figure 3**).

AID as a Polynucleotide Deaminase

As noticed when AID was first identified (39), its sequence reveals a domain common to cytidine deaminases, which includes a motif (His-X-Glu separated by a few amino acids from Pro-Cys-X-X-Cys) that is likely involved in zinc coordination. The sequence of the AID cytidine deaminase domain is most similar to that of the APOBEC proteins, with APOBEC1 in particular showing the closest homology. Indeed, the AID and APOBEC1

loci are closely linked in the genomes of man and mouse (21, 39, 69).

The AID/APOBEC family members appear to act specifically on cytidine in the context of polynucleotides; they do not deaminate free cytosine or its corresponding nucleoside/nucleotides (56, 60, 70). In vitro studies reveal that purified AID acts specifically on single- (as opposed to double-) stranded DNA substrates (47, 54–59); it has not been found to deaminate RNA (54, 56). Interestingly, APOBEC1 (while functioning physiologically as an RNA deaminase) can deaminate DNA in vitro (70, 71), although it does not substitute for AID in functional antibody diversification assays (72, 73).

Several models have been proposed for the three-dimensional structures of APOBEC-family deaminases on the basis of known mononucleotide deaminase structures (74–77), but the models differ from each another and remain speculative. However, the models coincide in predicting that AID, like all known cytidine deaminases, might have a quaternary structure. Indeed, there is experimental evidence suggesting that AID forms functional homodimers, and the region in between amino acids Thr27 and His56 has been defined as essential for AID dimerization (78) (**Figure 3**). Interestingly, native AID purified from rabbit appendix shows two forms differing in apparent electrophoretic mobility, compatible with the presence of both monomer and homodimer (79). Intriguingly, these dimers seem to be resistant to conventional denaturing conditions, a feature shared by the oligomers formed by recombinant AID (56).

Subcellular Localization of AID

As might be expected for an enzyme that can mutate DNA, the access of AID to the nucleus is regulated. The N-terminal region of AID has a high net positive charge and contains what may well be a bipartite nuclear localization signal (80). Thus, this region of AID is sufficient to mediate nuclear compart-

mentalization when fused to green fluorescent protein (GFP). Furthermore, AID-GFP chimaeras carrying deletions of this N-terminal region of AID fail to accumulate in the nucleus under conditions when nuclear export is inhibited (80). However, the N-terminal region of AID alone is not sufficient to mediate nuclear import of a larger fusion protein (81), so there might be a need for additional factors to cooperate in sending AID to the nucleus.

The C-terminal portion of AID contains a strong nuclear export signal, which explains why most AID protein is found in the cytoplasm (65, 80–82). The significance of this ~30-amino acid long C-terminal region was first appreciated when Ta et al. (74) showed the presence of mutations within this region in human patients with hyper-IgM syndromes. This region of AID is absolutely required for CSR but seems to be dispensable for both SHM and gene conversion (74, 83, 84). Thus, it is tempting to speculate that the C-terminal region of AID engages in interactions that are essential for CSR as opposed to SHM. Although MDM2 (a protein that controls p53) and DNA-PKcs (a protein normally implicated in the joining of broken DNA ends) have been found to be able to interact with the C-terminal portion of AID in either yeast two-hybrid or pull-down assays (85, 86), at present it is not known whether such interactions are of physiological significance. Thus, it is not yet possible to conclude whether the importance of the C-terminal portion of AID for the execution of CSR does indeed reflect any CSR-specific protein interactions of AID.

Interestingly, dimerization of AID has also recently been postulated as essential for CSR because mutations within the dimerization region of AID fail to reconstitute during in vitro CSR assays (78). This is consistent with the suggestion that AID may act in CSR through simultaneous deamination of overlapping WRC motifs on opposite DNA strands (60). However, although the C terminus of AID is leucine rich (reminiscent of protein-protein interaction domains), this

region of AID has not been shown to partake in oligomerization (74, 78).

Partners of AID

The physiological targeting of AID to the Ig locus might depend on proteins that interact specifically with AID. Indeed, the distinct recruitment of AID to sites of CSR or SHM suggests that the two processes might use separate factors for AID recruitment. However, so far, only two proteins have been shown to interact with AID in B cells: the replication protein A (RPA) (87) and the protein kinase A (PKA) alpha regulatory subunit (PKAr1 α) (88–90).

The interaction of AID with RPA could well provide insight into the mechanism by which AID accesses a single-stranded DNA substrate in vivo. This interaction dramatically increases the efficiency of AID-catalyzed deamination of plasmid substrates in vitro (87), possibly by stabilizing single-stranded DNA during transcription and allowing AID access to it. Thus, RPA binds specifically to single-stranded DNA, whereas the interaction with PKAr1 α could reflect a pathway for regulation of AID activity. AID associates with PKAr1 α in the cytoplasm; this interaction allows PKA to phosphorylate AID on two residues within its N terminus (Ser27 and Thr38 in human AID) (**Figure 3**) (88–90). Although nonphosphorylated AID is still enzymatically competent, this posttranslational modification is important for the biological activity of the enzyme. AID with mutated PKA-phosphorylation sites is much less potent (although not completely inactive) in triggering CSR (88–90) and SHM (89). Furthermore, the observations that mice with constitutively activated PKA display enhanced CSR and that PKA inhibitors prevent in vitro class switching provide compelling evidence for a physiological regulatory role of PKA in antibody diversification (90). However, the molecular basis for the importance of phosphorylation is not yet fully defined. Phosphorylated AID is not especially enriched in

the nucleus, although it is preferentially associated with chromatin (89). Because phosphorylation by PKA affects the interaction of AID with the 32-kDa subunit of RPA (87, 88), it is possible that PKA influences the targeting of AID to single-stranded DNA to potentiate antibody gene diversification.

TARGETING OF AID AND THE LINKAGE TO TRANSCRIPTION

The mechanism by which AID is specifically targeted to the IgV gene for the performance of SHM and the way in which this IgV target is rendered single stranded and accessible to AID action are areas of considerable ignorance. Early studies using a variety of transgenic and knockin mice revealed that the transcription-regulatory elements of the Ig locus play a major role in the recruitment of hypermutation [reviewed in (17, 91)]. Thus, the enhancers (but not the body of the IgV gene itself) play a significant role in mutation recruitment, and the IgV gene promoter defines the 5' border of the mutation domain, although it can be replaced by a heterologous promoter. The molecular mechanism by which these transcription regulatory elements function in potentiating SHM is, however, not clearly defined.

These observations would obviously be consistent with a role for transcription in the recruitment of AID for SHM, and indeed, transcription enhances in vitro AID-mediated deamination of synthetic substrates as well as AID-catalyzed mutation of *Escherichia coli* (47, 54, 55, 57, 58, 60, 92). There is also a correlation between AID recruitment and transcription of the switch region target in the case of CSR (93, 94). Transfected AID has been reported to immunoprecipitate with RNA polymerase II from in vitro-activated B cells, suggesting a physical interaction between AID and the transcription complex (93). However, the nature of the molecular association between AID recruitment and target gene transcription clearly needs to be defined in greater detail.

RPA: replication protein A

PKA: protein kinase A

Analysis of the patterns of nucleotide substitution introduced into IgV genes during SHM suggests that, *in vivo*, AID deaminates dC residues on the transcribed or nontranscribed DNA strand with equal probability [reviewed in (17) and discussed in (64)]. This contrasts with many studies of AID-catalyzed deamination of DNA targets transcribed *in vitro* (55, 57–59), although deamination of both DNA strands has been observed when supercoiled plasmids were used as *in vitro* substrates (95), as well as with certain loss-of-function reporter selections (92), suggesting that AID is inherently able to target both transcribed and nontranscribed strands.

Even though mutations accumulate during SHM, with only a very small number (e.g., one to five) of unclustered mutations incorporated during each round of cell division [reviewed in (17)], recombinant AID frequently appears to act processively *in vitro*, generating multiple linked deaminations on the same DNA strand (57). It may be that there are factors *in vivo* that prevent AID from acting processively or, alternatively, that AID does indeed act processively *in vivo*, but many of the U:G mutations generated are actually repaired such that multiple linked mutations are not actually fixed *in vivo* (94).

The Ig loci constitute the major physiological targets of AID-mediated DNA deamination, yet other targets can be subject to deamination, especially under conditions of AID overexpression (44, 45, 96–99). Most genes that attract the SHM process contain binding sites for the E2A transcription factors (100, 101), and furthermore, antibody diversification in chicken DT40 cells is affected by modulating E2A levels (102, 103). However, demonstration of a direct role of E2A factors in the execution of SHM is lacking. Whatever the mechanism, it appears easy to perturb the proper targeting of AID. Thus, overexpression of AID or use of transfected (as opposed to endogenous) DNA targets can yield high levels of mutation accumulation in nonphysiological target sequences (44, 45, 96–

99). This can complicate experimental strategies for investigating the physiological targeting mechanisms. Nevertheless, it is apparent that targeting of AID to nonimmunoglobulin targets is of pathological concern because it is a feature of many B cell tumors where it likely underpins the chromosomal translocations and proto-oncogene mutations found in many Burkitt and follicular B cell lymphomas (104).

PROCESSING THE AID-GENERATED U:G LESION

The diversity of changes that result from AID-triggered deamination (the different types of nucleotide substitutions introduced during SHM as well as the triggering of Ig gene conversion and CSR) reflect the different pathways used for processing the initiating U:G lesion. If the U:G lesion generated by AID-mediated deamination is simply replicated, then the nucleotide substitutions generated during SHM are restricted to C to T and G to A transition substitutions (**Figure 2**). All other types of nucleotide substitutions (transversions at C:G pairs as well as all substitutions at A:T pairs) depend on recognition of the U:G lesion by proteins that are normally involved in either base excision or mismatch repair pathways. The question obviously arises as to how they are exploited in SHM to assist antibody diversification rather than performing their usual function of error-free repair.

Uracil Excision by UNG

The most commonly used pathway to remove uracil from DNA is base excision repair. This is initiated by enzymes that excise uracil from the deoxyribose-phosphate backbone. Four such uracil-DNA glycosylases have been described in vertebrates: UNG, SMUG1, TDG, and MBD4 (105). UNG provides the main uracil-excision activity measured in cell extracts and has also proven fundamental in Ig diversification wherein it mediates the major

pathways, leading to transversion mutations at G:C base pairs (61, 106), Ig gene conversion (62, 66), and CSR (63, 106). The importance of UNG is revealed by the profound effect that UNG deficiency has on antibody gene diversification (62, 63, 66, 106). None of the other uracil-DNA glycosylases normally act as a backup for UNG in SHM or CSR (64, 106–108), suggesting that it is not only uracil excision itself that is important but specifically excision by UNG. This notion is reinforced when transgenic overexpression of SMUG1 partially substitutes for UNG when UNG is absent, yet SMUG1 overexpression in normal B cells actually diminishes the frequency of Ig diversification, probably by favoring faithful repair of the uracil instead of diversification (108). The fact that UNG, rather than the other uracil-excision enzymes, functions in antibody diversification, it is therefore probably not just a reflection on the abundance of UNG but also likely a correlation with its expression regulation during the cell cycle as well as with its association with replication foci and other proteins (105, 108, 109).

Recognition of the U:G Mismatch by MSH2/MSH6

In addition to generating a foreign base in DNA that can be recognized by the excision enzymes of the base excision repair pathway, introduction of uracil into DNA through AID-catalyzed deamination of deoxycytidine also creates a mismatch because uracil is placed opposite guanine. This U:G mispair can be recognized by the MSH2/MSH6 mismatch recognition heterodimer (68). Such recognition can trigger a patch DNA synthesis process in which additional mutations are introduced (the second phase of SHM), leading to nucleotide substitutions at A:T pairs (see below). In the absence of MSH2 or MSH6, the accumulation of mutations at A:T pairs is substantially diminished (31–33, 67, 110–112).

UNG and MSH2/MSH6 Provide Alternative Pathways for Processing the U:G Lesion

Although mutation accumulation at A:T pairs is diminished to a large extent in mice deficient in MSH2/MSH6, mutations at A:T are entirely ablated in mice that are simultaneously deficient in both MSH2 and UNG. Indeed, CSR is similarly ablated by a UNG/MSH2-double deficiency and the grosser genomic changes (e.g., deletions/insertions) that occur as occasional by-products of DNA lesion repair are also reduced (64). In such mice, AID-triggered DNA deamination leads solely to the generation of C to T and G to A transition substitutions at C:G pairs [this applies to both the IgV and switch region targets (64)]. Thus, it appears that in the absence of both UNG and MSH2, the U:G mispair is scarcely recognized as a lesion and is simply replicated.

It is not known whether the UNG and MSH2/MSH6 pathways compete for the recognition of uracil or act on uracils produced at different times during the cell cycle. Some degree of competition in recognizing the same lesion must occur because uracil excision can provide a minor backup pathway for the second phase of SHM and because mismatch recognition provides a backup pathway for isotype class switching (64, 108). However, it could be that there is a more complex interplay between the two pathways. A deficiency in components of the mismatch recognition/processing pathways affects mostly the second phase of SHM (mutations at A:T pairs), but it also results in a significant increase in the proportion of substitutions at C:G pairs that are transitions as opposed to transversions (31, 33, 67, 110–112). This as yet unexplained observation suggests that a deficiency in these factors results in diminished uracil excision with the data pointing to an uncharacterized interaction between the mismatch-recognition and uracil-excision pathways.

Transversion

mutation: a base change in DNA in which a pyrimidine (C or T) is replaced by a purine (A or G), or a purine is replaced by a pyrimidine

Uracil excision:

removal of the uracil base from deoxyuridine in DNA through cleavage of the glycosidic bond by a uracil-DNA glycosylase enzyme, resulting in an abasic site

Replication Over the UNG-Generated Abasic Site

The abasic sites generated by UNG-mediated excision of the AID-generated uracil constitute lesions that block (or, at least slow down) DNA replication catalyzed by conventional, replicative DNA polymerases. Several non-processive DNA polymerases have been characterized in recent years, many of which might play a role in synthesis across DNA lesions (lesion bypass synthesis) (113). Several such DNA polymerases have been suggested as participants in SHM (114).

There is good evidence of a role for the deoxycytidyl transferase Rev1 in DNA synthesis across abasic sites in the Ig genes. Not only is Rev1 effective in synthesizing DNA across such templates, but Rev1 also has a strong preference for inserting deoxycytidine residues in the newly synthesized strand, which would result in C to G and G to C transversion mutations (113, 115). Such transversion mutations are significantly diminished during IgV hypermutation in chicken DT40 B cells as well as in mouse B cells that have been rendered deficient in Rev1 (116, 117). However, Rev1 as well as many other translesion polymerases are expected to insert only a single base opposite the abasic site; extension would require the subsequent action of a mismatch-extending polymerase (113, 115). This might explain the perturbation of SHM reported following DNA polymerase ζ knockdown (118, 119).

It is very likely that translesion polymerases other than Rev1 also play a role in bypassing the abasic site. However, the interpretation of results is complicated by the differences between systems analyzed and the likely redundancy of these polymerases, whereby deficiency in one polymerase likely results in a compensatory action by another one, a circumstance that might not be apparent from the ensuing mutation pattern. For instance, a deficiency in DNA polymerase θ results in a relatively modest effect on mutation accumulation, but two groups obtained

different results regarding whether it affects all mutations similarly or preferentially affects mutations at C:G pairs (120–122). Similarly, a deficiency in DNA polymerase ι correlates with diminished mutation accumulation in a B cell line, but this polymerase is dispensable for SHM in knockout mice (123, 124).

INTRODUCTION OF MUTATIONS AT A:T PAIRS

All of the nucleotide substitutions occurring at C:G pairs during SHM can be accounted for by faithful replication over the AID-generated U:G lesion together with replication over the UNG-generated abasic site (where the polymerase must display some initiative regarding base insertion opposite such a non-instructional lesion). However, an additional mutagenic process must be invoked to explain the accumulation of mutations at A:T pairs that occurs during SHM.

This second phase of mutation creation, as discussed above, appears to occur during MSH2/MSH6-triggered patch repair of the initiating U:G lesion in a process that also involves exonuclease I (31–33, 125). There is clear evidence that DNA polymerase η plays a critical role in this process, with both humans and mice deficient in this polymerase exhibiting a striking reduction in mutation accumulation at A:T (but not C:G) pairs during SHM (126–128).

Relationship to Conventional Mismatch Repair

Although MSH2/MSH6 and exonuclease I are components of the postreplicative mismatch repair pathway, it is unlikely for several reasons that the second phase of SHM occurs as part of conventional postreplicative mismatch repair. First, the second phase of SHM generates mutations at A:T pairs, and there is no evidence that such mutations are a normal part of postreplicative mismatch repair. Second, conventional mismatch repair takes place soon after DNA replication;

the AID-generated U:G mismatch would already have been resolved by DNA replication prior to classical mismatch repair unless AID deaminates the IgV gene at the same time as or immediately following the passage of the replication fork. Third, whereas postreplicative mismatch repair is dependent on PMS2, MLH1, and MLH3, deficiency in any of these three proteins has little, or at most a minor, effect on the second phase of SHM (32, 129–134). Fourth, although the DNA polymerases involved in conventional mismatch repair have not been unambiguously defined, DNA polymerase δ is currently the favored candidate with no evidence for a dependence on DNA polymerase η [reviewed in (135)]. In contrast, DNA polymerase η clearly plays a major role in the introduction of A:T pairs during SHM (126–128).

Mechanism of Mutagenesis at A:T Pairs

How exactly the mutations at A:T are generated during this process is still unclear. The dependence on DNA polymerase η suggests that this polymerase plays a critical role in the incorporation of the mutations—a suggestion that is nicely supported by the recent finding that MSH2/MSH6 both interacts with and stimulates DNA polymerase η (68). The most straightforward explanation for mutation is that DNA polymerase η is recruited to resynthesize the DNA patch that has been degraded by exonuclease I and that the mutations simply reflect the relatively high error frequency and misincorporation preferences, which have been described for polymerase η from *in vitro* studies (68, 126–128, 136, 137). However, although it is certainly simple and attractive to propose that the mutation generation during this phase of SHM occurs through base mispairing as a consequence of untemplated polymerase error, this leaves several issues unresolved, and possible alternative mechanisms (such as misincorporation of a noncanonical base) cannot be definitively excluded [discussed in (138)].

With regard to the nature of a patch synthesized during the second phase of SHM, considerations of the linkage of C:G and A:T mutations as well as of the span of the IgV C:G and A:T mutation domains give a pointer as to the length of this patch (around 30–100 nucleotides). Evidence from several studies of DNA strand polarity with respect to the accumulation of mutations at A:T (but not C:G) pairs during SHM also suggests that the length of the mutagenic repair patch is likely to differ on the two DNA strands, possibly reflecting linkage to transcription or passage of the DNA replication fork. These issues are discussed more fully elsewhere (138).

There is clearly much to learn about the molecular processes involved in the second phase of SHM; a better understanding of the DNA mismatch repair pathways and of the role of polymerase η in somatic hypermutation may well prove to be parallel roads.

MUTATION VERSUS REPAIR

With the exception of AID itself, our current understanding of SHM invokes the actions of proteins that are widely distributed and that normally fulfill a role in DNA repair. The question therefore arises as to how the normal function of these proteins in predominantly error-free repair is subverted so as to achieve mutagenesis. This applies to both phases of mutation creation.

Thus, with regard to the first phase, why is the abasic site that is generated by UNG-mediated uracil excision used as a template for DNA synthesis (leading to mutation) rather than excised and replaced by a cytidine residue, using the guanine on the other strand as template, as would occur with conventional error-free base excision repair? We have previously proposed that coordination of the UNG-mediated uracil-excision step with passage of the DNA replication fork may play a critical role in ensuring a mutagenic outcome (108). Thus, if excision of the uracil occurs just before passage of the replication fork without allowing repair of the

abasic site, then the abasic site could provide a noninstructional lesion on the template strand for DNA synthesis during replication. However, although attractive and consistent with the known association of UNG with replication foci (109), there is no direct evidence of a link between mutation fixation and replication.

Other mechanisms to explain mutagenesis at the Ig locus have also been proposed. The Mre11-NBS-Rad50 protein complex was recently found to cleave abasic sites in single-stranded DNA (139), suggesting a way of creating single strand breaks at the Ig locus following the sequential action of AID and UNG. Mutation has been proposed to then occur through error-prone repair of the break, but the mechanism has not been defined in any detail (139, 140).

With regard to the second phase of mutation, the MSH2/MSH6-triggered patch repair must be highly mutagenic in order to yield the abundance of mutations at A:T pairs that is observed during SHM. Assuming that the patch DNA synthesis associated with conventional mismatch repair does not show this striking infidelity, then some explanation is required to account for the very high mutation frequency associated with the second phase of SHM.

CONCLUSION

It has been 50 years since Burnet proposed the clonal selection theory of antibody formation, which required an unprecedented genetic process for generating antibody diversity. We now know that there are in fact two key modifications of the immunoglobulin loci that underpin the generation of antibody gene diversity. The first is site-specific gene rearrangement (catalyzed by the RAG1/RAG2 recombinase and involving terminal deoxynucleotidyl transferase) and yields a primary repertoire of functional antibody genes. The second is targeted deamination of deoxycytidine residues in the Ig loci (catalyzed by AID), which generates a lesion that is re-

solved by ubiquitous DNA repair enzymes to yield a diversity of nucleotide substitutions across the IgV (somatic hypermutation). In other circumstances, the same process of AID-triggered DNA deamination is also used as a trigger for IgV gene conversion (in chickens) or in IgC CSR (in many vertebrates).

RAG-mediated rearrangement and AID-mediated deamination both appear to have arisen around the time of vertebrate evolution. However, in man and mouse, RAG-mediated recombination is now used to generate the primary repertoire of antibody specificities, whereas AID-mediated deamination is used to enlarge this repertoire and drives the iterative process of mutation/selection that occurs following an antigen encounter and that underpins antibody affinity maturation.

Although the initial trigger (DNA deamination) and broad outlines of the molecular mechanism of SHM have been identified, there are many questions that still need to be answered. In particular, what is the means by which AID is preferentially targeted to the immunoglobulin loci? And is the second phase of SHM (in which mutations are introduced at A:T pairs) caused by erroneous base mispairing [essentially as envisioned in the Brenner-Milstein 1966 model (9)]? Or is there another mechanism at play?

Investigations of SHM have not only provided insight into the molecular basis of antibody gene diversification, they have also unveiled a previously unsuspected physiological mechanism for genetic change: the programmed alteration of the DNA coding information through targeted base modification. Interestingly, targeted deamination of cytosine in DNA appears not to be restricted to the adaptive immune system. A similar process is catalyzed by the AID-related APOBEC3 deaminases, a family of restriction factors that target deoxycytidines in lentiviral replication intermediates [reviewed in (141)]. Thus, although the presence of uracil in DNA probably constitutes one of the most primitive DNA lesions (142), the immune systems of

higher organisms appear to have evolved a strategy of intentionally generating targeted uracil lesions in either their own or in invading genomes, thereby providing a trigger for gene

diversification (the antibodies of the adaptive immune system) or for inactivation of foreign genes (the antileviral pathway in the innate immune system).

SUMMARY POINTS

1. Targeted deamination of deoxycytidine residues within immunoglobulin variable (IgV) genes generates a deoxyuridine lesion that provides a trigger for IgV gene diversification and antibody affinity maturation.
2. The targeted deoxycytidine deamination is catalyzed by activation-induced deaminase (AID), there being an association between transcription of the IgV gene target and AID recruitment.
3. AID is regulated by protein-protein interactions and posttranslational modifications, which likely ensure its specific action at the immunoglobulin loci, but our understanding of this area remains limited.
4. The deoxyuridine lesion resulting from AID-catalyzed deamination of deoxycytidine is detected by either the uracil-DNA glycosylase UNG or by the mismatch recognition heterodimer MSH2/MSH6 (as deamination of a C:G pair generates a U:G mismatch). Although both UNG- and MSH2/MSH6-mediated recognition would normally result in error-free DNA repair, in antibody diversification such recognition triggers mutagenic resolution.
5. Uracil excision by UNG mediates the major pathways of antibody diversification through somatic hypermutation at C:G pairs, IgV gene conversion, and isotype class switching. Recognition by MSH2/MSH6 provides only a partial backup for these processes in the absence of UNG.
6. Recognition by MSH2/MSH6 does, however, play a dominant role in triggering the mutagenic patch repair that leads to the introduction of mutations at A:T.
7. Two atypical DNA polymerases, Rev1 and DNA polymerase η , have been shown to have strong, nonredundant roles in antibody diversification with their ablation having an effect that is not compensated by other DNA polymerases.

FUTURE ISSUES

1. How is AID specifically targeted to the IgV gene and why can it be recruited to some but not other nonimmunoglobulin targets? In particular, what are the molecular details of the association between hypermutation and transcription?
2. When during the cell cycle does the generation and resolution of the U:G lesions occur and why are these lesions fixed as mutations rather than subjected to conventional repair?
3. What is the mechanism responsible for introducing mutations at A:T pairs and, in particular, does DNA polymerase η generate these mutations at A:T pairs through base mispairing owing to spontaneous polymerase error?
4. What are the three-dimensional structure and detailed catalytic features of AID?

ACKNOWLEDGMENTS

M.S.N. is indebted to the Royal Society for supporting a short visit to the laboratory of Dr. David Tarlinton at the WEHI, Melbourne, where part of this review was drafted.

LITERATURE CITED

1. Tonegawa S. 1983. *Nature* 302:575–81
2. Reynaud CA, Bertocci B, Dahan A, Weill JC. 1994. *Adv. Immunol.* 57:353–78
3. Maizels N. 2005. *Annu. Rev. Genet.* 39:23–46
4. Burnet FM. 1957. *Aust. J. Sci.* 20:67–69
5. Lederberg J. 1959. *Science* 129:1649–53
6. Hilschmann N, Craig LC. 1965. *Proc. Natl. Acad. Sci. USA* 53:1403–9
7. Milstein C. 1966. *Biochem. J.* 101:352–68
8. Putnam FW, Titani K, Whitley EJ. 1966. *Proc. R. Soc. London Ser. B* 166:124–37
9. Brenner S, Milstein C. 1966. *Nature* 211:242–43
10. Antibodies. 1967. *Cold Spring Harbor Symp. Quant. Biol.* 32:(Whole volume)
11. Weigert MG, Cesari IM, Yonkovich SJ, Cohn M. 1970. *Nature* 228:1045–47
12. Bernard O, Hozumi N, Tonegawa S. 1978. *Cell* 15:1133–44
13. O'Brien RL, Brinster RL, Storb U. 1987. *Nature* 326:405–9
14. Milstein C, Neuberger MS. 1996. *Adv. Protein Chem.* 49:451–85
15. Rajewsky K. 1996. *Nature* 381:751–58
16. 1998. *Immunol. Rev.* 162:(Whole volume)
17. Neuberger MS, Milstein C. 1995. *Curr. Opin. Immunol.* 7:248–54
18. Chaudhuri J, Alt FW. 2004. *Nat. Rev. Immunol.* 4:541–52
19. Reynaud CA, Mackay CR, Muller RG, Weill JC. 1991. *Cell* 64:995–1005
20. Hinds-Frey KR, Nishikata H, Litman RT, Litman GW. 1993. *J. Exp. Med.* 178:815–24
21. Conticello SG, Thomas CJ, Petersen-Mahrt SK, Neuberger MS. 2005. *Mol. Biol. Evol.* 22:367–77
22. Fugmann SD, Messier C, Novack LA, Cameron RA, Rast JP. 2006. *Proc. Natl. Acad. Sci. USA* 103:3728–33
23. Golding GB, Gearhart PJ, Glickman BW. 1987. *Genetics* 115:169–76
24. Milstein C, Neuberger MS, Staden R. 1998. *Proc. Natl. Acad. Sci. USA* 95:8791–94
25. Sharpe MJ, Milstein C, Jarvis JM, Neuberger MS. 1991. *EMBO J.* 10:2139–45
26. Rogozin IB, Kolchanov NA. 1992. *Biochim. Biophys. Acta* 1171:11–18
27. Betz AG, Rada C, Pannell R, Milstein C, Neuberger MS. 1993. *Proc. Natl. Acad. Sci. USA* 90:2385–88
28. Wagner SD, Milstein C, Neuberger MS. 1995. *Nature* 376:732
29. Jolly CJ, Wagner SD, Rada C, Klix N, Milstein C, Neuberger MS. 1996. *Semin. Immunol.* 8:159–68
30. Kepler TB. 1997. *Mol. Biol. Evol.* 14:637–43
31. Phung QH, Winter DB, Cranston A, Tarone RE, Bohr VA, et al. 1998. *J. Exp. Med.* 187:1745–51
32. Frey S, Bertocci B, Delbos F, Quint L, Weill JC, Reynaud CA. 1998. *Immunity* 9:127–34
33. Rada C, Ehrenstein MR, Neuberger MS, Milstein C. 1998. *Immunity* 9:135–41
34. Bachl J, Wabl M. 1996. *Proc. Natl. Acad. Sci. USA* 93:851–55
35. Denepoux S, Razanajaona D, Blanchard D, Meffre G, Capra JD, et al. 1997. *Immunity* 6:35–46
36. Sale JE, Neuberger MS. 1998. *Immunity* 9:859–69

37. Wilson M, Hsu E, Marcuz A, Courtet M, Du Pasquier L, Steinberg C. 1992. *EMBO J.* 11:4337–47
38. Neuberger MS, Harris RS, Di Noia J, Petersen-Mahrt SK. 2003. *Trends Biochem. Sci.* 28:305–12
39. Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, et al. 1999. *J. Biol. Chem.* 274:18470–76
40. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. 2000. *Cell* 102:553–63
41. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, et al. 2000. *Cell* 102:565–75
42. Arakawa H, Hauschild J, Buerstedde JM. 2002. *Science* 295:1301–6
43. Harris RS, Sale JE, Petersen-Mahrt SK, Neuberger MS. 2002. *Curr. Biol.* 12:435–38
44. Martin A, Bardwell PD, Woo CJ, Fan M, Shulman MJ, Scharff MD. 2002. *Nature* 415:802–6
45. Okazaki IM, Kinoshita K, Muramatsu M, Yoshikawa K, Honjo T. 2002. *Nature* 416:340–45
46. Blanc V, Davidson NO. 2003. *J. Biol. Chem.* 278:1395–98
47. Petersen-Mahrt SK, Harris RS, Neuberger MS. 2002. *Nature* 418:99–103
48. Ehrenstein MR, Neuberger MS. 1999. *EMBO J.* 18:3484–90
49. Neuberger MS, Lanoue A, Ehrenstein MR, Batista FD, Sale JE, Williams GT. 1999. *Cold Spring Harbor Symp. Quant. Biol.* 64:211–16
50. Maizels N. 1995. *Cell* 83:9–12
51. Weill JC, Reynaud CA. 1996. *Immunol. Today* 17:92–97
52. Sale JE, Calandrini DM, Takata M, Takeda S, Neuberger MS. 2001. *Nature* 412:921–26
53. Arakawa H, Saribasak H, Buerstedde JM. 2004. *PLoS Biol.* 2:E179
54. Bransteitter R, Pham P, Scharff MD, Goodman MF. 2003. *Proc. Natl. Acad. Sci. USA* 100:4102–7
55. Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW. 2003. *Nature* 422:726–30
56. Dickerson SK, Market E, Besmer E, Papavasiliou FN. 2003. *J. Exp. Med.* 197:1291–96
57. Pham P, Bransteitter R, Petruska J, Goodman MF. 2003. *Nature* 424:103–7
58. Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC. 2003. *Nat. Immunol.* 4:452–56
59. Sohail A, Klapacz J, Samaranayake M, Ullah A, Bhagwat AS. 2003. *Nucleic Acids Res.* 31:2990–94
60. Beale RC, Petersen-Mahrt SK, Watt IN, Harris RS, Rada C, Neuberger MS. 2004. *J. Mol. Biol.* 337:585–96
61. Di Noia J, Neuberger MS. 2002. *Nature* 419:43–48
62. Di Noia JM, Neuberger MS. 2004. *Eur. J. Immunol.* 34:504–8
63. Imai K, Slupphaug G, Lee WI, Revy P, Nonoyama S, et al. 2003. *Nat. Immunol.* 4:1023–28
64. Rada C, Di Noia JM, Neuberger MS. 2004. *Mol. Cell* 16:163–71
65. Rada C, Jarvis JM, Milstein C. 2002. *Proc. Natl. Acad. Sci. USA* 99:7003–8
66. Saribasak H, Saribasak NN, Ipek FM, Ellwart JW, Arakawa H, Buerstedde JM. 2006. *J. Immunol.* 176:365–71
67. Wiesendanger M, Kneitz B, Edelmann W, Scharff MD. 2000. *J. Exp. Med.* 191:579–84
68. Wilson TM, Vaisman A, Martomo SA, Sullivan P, Lan L, et al. 2005. *J. Exp. Med.* 201:637–45
69. Jarmuz A, Chester A, Bayliss J, Gisbourne J, Dunham I, et al. 2002. *Genomics* 79:285–96
70. Petersen-Mahrt SK, Neuberger MS. 2003. *J. Biol. Chem.* 278:19583–86
71. Harris RS, Petersen-Mahrt SK, Neuberger MS. 2002. *Mol. Cell* 10:1247–53

72. Fugmann SD, Rush JS, Schatz DG. 2004. *Eur. J. Immunol.* 34:844–49
73. Krause K, Marcu KB, Greeve J. 2006. *Mol. Immunol.* 43:295–307
74. Ta VT, Nagaoka H, Catalan N, Durandy A, Fischer A, et al. 2003. *Nat. Immunol.* 4:843–48
75. Zaim J, Kierzek AM. 2003. *Nat. Immunol.* 4:1153
76. Xie K, Sowden MP, Dance GS, Torelli AT, Smith HC, Wedekind JE. 2004. *Proc. Natl. Acad. Sci. USA* 101:8114–19
77. Huthoff H, Malim MH. 2005. *Virology* 334:147–53
78. Wang J, Shinkura R, Muramatsu M, Nagaoka H, Kinoshita K, Honjo T. 2006. *J. Biol. Chem.* 281:19115–23
79. Yang G, Obiakor H, Sinha RK, Newman BA, Hood BL, et al. 2005. *Proc. Natl. Acad. Sci. USA* 102:17083–88
80. Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, et al. 2004. *Proc. Natl. Acad. Sci. USA* 101:1975–80
81. McBride KM, Barreto V, Ramiro AR, Stavropoulos P, Nussenzweig MC. 2004. *J. Exp. Med.* 199:1235–44
82. Brar SS, Watson M, Diaz M. 2004. *J. Biol. Chem.* 279:26395–401
83. Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM, Nussenzweig MC. 2003. *Mol. Cell* 12:501–8
84. Shinkura R, Ito S, Begum NA, Nagaoka H, Muramatsu M, et al. 2004. *Nat. Immunol.* 5:707–12
85. MacDuff DA, Neuberger MS, Harris RS. 2006. *Mol. Immunol.* 43:1099–108
86. Wu X, Geraldles P, Platt JL, Cascalho M. 2005. *J. Immunol.* 174:934–41
87. Chaudhuri J, Khuong C, Alt FW. 2004. *Nature* 430:992–98
88. Basu U, Chaudhuri J, Alpert C, Dutt S, Ranganath S, et al. 2005. *Nature* 438:508–11
89. McBride KM, Gazumyan A, Woo EM, Barreto VM, Robbiani DF, et al. 2006. *Proc. Natl. Acad. Sci. USA* 103:8798–803
90. Pasqualucci L, Kitaura Y, Gu H, Dalla-Favera R. 2006. *Proc. Natl. Acad. Sci. USA* 103:395–400
91. Odegard VH, Schatz DG. 2006. *Nat. Rev. Immunol.* 6:573–83
92. Besmer E, Market E, Papavasiliou FN. 2006. *Mol. Cell. Biol.* 26:4378–85
93. Nambu Y, Sugai M, Gonda H, Lee CG, Katakai T, et al. 2003. *Science* 302:2137–40
94. Xue K, Rada C, Neuberger MS. 2006. *J. Exp. Med.* 203:2085–94
95. Shen HM, Storb U. 2004. *Proc. Natl. Acad. Sci. USA* 101:12997–3002
96. Okazaki IM, Hiai H, Kakazu N, Yamada S, Muramatsu M, et al. 2003. *J. Exp. Med.* 197:1173–81
97. Parsa JY, Basit W, Wang CL, Gommerman JL, Carlyle JR, Martin A. 2006. *Mol. Immunol.* 44:567–75
98. Wang CL, Harper RA, Wabl M. 2004. *Proc. Natl. Acad. Sci. USA* 101:7352–56
99. Martin A, Scharff MD. 2002. *Proc. Natl. Acad. Sci. USA* 99:12304–8
100. Michael N, Shen HM, Longerich S, Kim N, Longacre A, Storb U. 2003. *Immunity* 19:235–42
101. Kotani A, Okazaki IM, Muramatsu M, Kinoshita K, Begum NA, et al. 2005. *Proc. Natl. Acad. Sci. USA* 102:4506–11
102. Schoetz U, Cervelli M, Wang YD, Fiedler P, Buerstedde JM. 2006. *J. Immunol.* 177:395–400
103. Conlon TM, Meyer KB. 2006. *Eur. J. Immunol.* 36:139–48
104. de Yebenes VG, Ramiro AR. 2006. *Trends Mol. Med.* 12:432–39
105. Krokan HE, Drablos F, Slupphaug G. 2002. *Oncogene* 21:8935–48

106. Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS. 2002. *Curr. Biol.* 12:1748–55
107. Bardwell PD, Martin A, Wong E, Li Z, Edelmann W, Scharff MD. 2003. *J. Immunol.* 170:1620–24
108. Di Noia JM, Rada C, Neuberger MS. 2006. *EMBO J.* 25:585–95
109. Otterlei M, Warbrick E, Nagelhus TA, Haug T, Slupphaug G, et al. 1999. *EMBO J.* 18:3834–44
110. Bertocci B, Quint L, Delbos F, Garcia C, Reynaud CA, Weill JC. 1998. *Immunity* 9:257–65
111. Li Z, Scherer SJ, Ronai D, Iglesias-Ussel MD, Peled JU, et al. 2004. *J. Exp. Med.* 200:47–59
112. Martomo SA, Yang WW, Gearhart PJ. 2004. *J. Exp. Med.* 200:61–68
113. Prakash S, Johnson RE, Prakash L. 2005. *Annu. Rev. Biochem.* 74:317–53
114. Seki M, Gearhart PJ, Wood RD. 2005. *EMBO Rep.* 6:1143–48
115. Nelson JR, Lawrence CW, Hinkle DC. 1996. *Nature* 382:729–31
116. Ross AL, Sale JE. 2006. *Mol. Immunol.* 43:1587–94
117. Jansen JG, Langerak P, Tsaalbi-Shtylik A, van den Berk P, Jacobs H, de Wind N. 2006. *J. Exp. Med.* 203:319–23
118. Diaz M, Verkoczy LK, Flajnik MF, Klinman NR. 2001. *J. Immunol.* 167:327–35
119. Zan H, Komori A, Li Z, Cerutti A, Schaffer A, et al. 2001. *Immunity* 14:643–53
120. Masuda K, Ouchida R, Hikida M, Nakayama M, Ohara O, et al. 2006. *DNA Repair.* 5:1384–91
121. Masuda K, Ouchida R, Takeuchi A, Saito T, Koseki H, et al. 2005. *Proc. Natl. Acad. Sci. USA* 102:13986–91
122. Zan H, Shima N, Xu Z, Al-Qahtani A, Evinger Iii AJ, et al. 2005. *EMBO J.* 24:3757–69
123. Faili A, Aoufouchi S, Flatter E, Gueranger Q, Reynaud CA, Weill JC. 2002. *Nature* 419:944–47
124. Martomo SA, Yang WW, Vaisman A, Maas A, Yokoi M, et al. 2006. *DNA Repair* 5:392–98
125. Bardwell PD, Woo CJ, Wei K, Li Z, Martin A, et al. 2004. *Nat. Immunol.* 5:224–29
126. Delbos F, De Smet A, Faili A, Aoufouchi S, Weill JC, Reynaud CA. 2005. *J. Exp. Med.* 201:1191–96
127. Martomo SA, Yang WW, Wersto RP, Ohkumo T, Kondo Y, et al. 2005. *Proc. Natl. Acad. Sci. USA* 102:8656–61
128. Zeng X, Winter DB, Kasmer C, Kraemer KH, Lehmann AR, Gearhart PJ. 2001. *Nat. Immunol.* 2:537–41
129. Ehrenstein MR, Rada C, Jones AM, Milstein C, Neuberger MS. 2001. *Proc. Natl. Acad. Sci. USA* 98:14553–58
130. Wu X, Tsai CY, Patam MB, Zan H, Chen JP, et al. 2006. *J. Immunol.* 176:5426–37
131. Li Z, Peled JU, Zhao C, Svetlanov A, Ronai D, et al. 2006. *DNA Repair* 5:675–82
132. Kim N, Bozek G, Lo JC, Storb U. 1999. *J. Exp. Med.* 190:21–30
133. Phung QH, Winter DB, Alrefai R, Gearhart PJ. 1999. *J. Immunol.* 162:3121–24
134. Kong Q, Maizels N. 1999. *Mol. Immunol.* 36:83–91
135. Kunkel TA, Erie DA. 2005. *Annu. Rev. Biochem.* 74:681–710
136. Pavlov YI, Rogozin IB, Galkin AP, Aksenova AY, Hanaoka F, et al. 2002. *Proc. Natl. Acad. Sci. USA* 99:9954–59
137. Rogozin IB, Pavlov YI, Bebenek K, Matsuda T, Kunkel TA. 2001. *Nat. Immunol.* 2:530–36
138. Neuberger MS, Di Noia JM, Beale RC, Williams GT, Yang Z, Rada C. 2005. *Nat. Rev. Immunol.* 5:171–78

139. Larson ED, Cummings WJ, Bednarski DW, Maizels N. 2005. *Mol. Cell* 20:367–75
 140. Yabuki M, Fujii MM, Maizels N. 2005. *Nat. Immunol.* 6:730–36
 141. Harris RS, Liddament MT. 2004. *Nat. Rev. Immunol.* 4:868–77
 142. Poole A, Penny D, Sjoberg BM. 2001. *Nat. Rev. Mol. Cell Biol.* 2:147–51
-

RELATED REVIEW

Honjo T, Kinoshita K, Muramatsu M. 2002. *Annu. Rev. Immunol.* 20:165–96



Contents

Mitochondrial Theme

The Magic Garden <i>Gottfried Schatz</i>	673
DNA Replication and Transcription in Mammalian Mitochondria <i>Maria Falkenberg, Nils-Göran Larsson, and Claes M. Gustafsson</i>	679
Mitochondrial-Nuclear Communications <i>Michael T. Ryan and Nicholas J. Hoogenraad</i>	701
Translocation of Proteins into Mitochondria <i>Walter Neupert and Johannes M. Herrmann</i>	723
The Machines that Divide and Fuse Mitochondria <i>Suzanne Hoppins, Laura Lackner, and Jodi Nunnari</i>	751
Why Do We Still Have a Maternally Inherited Mitochondrial DNA? Insights from Evolutionary Medicine <i>Douglas C. Wallace</i>	781
Molecular Mechanisms of Antibody Somatic Hypermutation <i>Javier M. Di Noia and Michael S. Neuberger</i>	1
Structure and Mechanism of Helicases and Nucleic Acid Translocases <i>Martin R. Singleton, Mark S. Dillingham, and Dale B. Wigley</i>	23
The Nonsense-Mediated Decay RNA Surveillance Pathway <i>Yao-Fu Chang, J. Saadi Imam, Miles F. Wilkinson</i>	51
Functions of Site-Specific Histone Acetylation and Deacetylation <i>Mona D. Shabbazian and Michael Grunstein</i>	75
The tmRNA System for Translational Surveillance and Ribosome Rescue <i>Sean D. Moore and Robert T. Sauer</i>	101
Membrane Protein Structure: Prediction versus Reality <i>Arne Elofsson and Gunnar von Heijne</i>	125

Structure and Function of Toll Receptors and Their Ligands <i>Nicholas J. Gay and Monique Gangloff</i>	141
The Role of Mass Spectrometry in Structure Elucidation of Dynamic Protein Complexes <i>Michal Sharon and Carol V. Robinson</i>	167
Structure and Mechanism of the 6-Deoxyerythronolide B Synthase <i>Chaitan Khosla, Yinyan Tang, Alice Y. Chen, Nathan A. Schnarr, and David E. Cane</i>	195
The Biochemistry of Methane Oxidation <i>Amanda S. Hakemian and Amy C. Rosenzweig</i>	223
Anthrax Toxin: Receptor Binding, Internalization, Pore Formation, and Translocation <i>John A.T. Young and R. John Collier</i>	243
Synapses: Sites of Cell Recognition, Adhesion, and Functional Specification <i>Soichiro Yamada and W. James Nelson</i>	267
Lipid A Modification Systems in Gram-negative Bacteria <i>Christian R.H. Raetz, C. Michael Reynolds, M. Stephen Trent, and Russell E. Bishop</i>	295
Chemical Evolution as a Tool for Molecular Discovery <i>S. Jarrett Wrenn and Pebr B. Harbury</i>	331
Molecular Mechanisms of Magnetosome Formation <i>Arash Komeili</i>	351
Modulation of the Ryanodine Receptor and Intracellular Calcium <i>Ran Zalk, Stephan E. Lebnart, and Andrew R. Marks</i>	367
TRP Channels <i>Kartik Venkatachalam and Craig Montell</i>	387
Studying Individual Events in Biology <i>Stefan Wennmalm and Sanford M. Simon</i>	419
Signaling Pathways Downstream of Pattern-Recognition Receptors and Their Cross Talk <i>Myeong Sup Lee and Young-Joon Kim</i>	447
Biochemistry and Physiology of Cyclic Nucleotide Phosphodiesterases: Essential Components in Cyclic Nucleotide Signaling <i>Marco Conti and Joseph Beavo</i>	481
The Eyes Absent Family of Phosphotyrosine Phosphatases: Properties and Roles in Developmental Regulation of Transcription <i>Jennifer Jemc and Ilaria Rebay</i>	513

Assembly Dynamics of the Bacterial MinCDE System and Spatial Regulation of the Z Ring <i>Joe Lutkenhaus</i>	539
Structures and Functions of Yeast Kinetochore Complexes <i>Stefan Westermann, David G. Drubin, and Georjana Barnes</i>	563
Mechanism and Function of Formins in the Control of Actin Assembly <i>Bruce L. Goode and Michael J. Eck</i>	593
Unsolved Mysteries in Membrane Traffic <i>Suzanne R. Pfeffer</i>	629
Structural Biology of Nucleocytoplasmic Transport <i>Atlanta Cook, Fulvia Bono, Martin Jinek, and Elena Conti</i>	647
The Postsynaptic Architecture of Excitatory Synapses: A More Quantitative View <i>Morgan Sheng and Casper C. Hoogenraad</i>	823

Indexes

Cumulative Index of Contributing Authors, Volumes 72–76	849
Cumulative Index of Chapter Titles, Volumes 72–76	853

Errata

An online log of corrections to *Annual Review of Biochemistry* chapters (if any, 1997 to the present) may be found at <http://biochem.annualreviews.org/errata.shtml>