Structural Phylogenetic Analysis of Activation-Induced Deaminase Function

H. Travis Ichikawa, Mark P. Sowden, Andrew T. Torelli, Jürgen Bachl, Pinwei Huang, Geoffrey S. C. Dance, Shauna H. Travis Ichikawa, Mark P. Sowden, Andrew T. Torelli, Jürgen Bachl, Pinwei Huang, Geoffrey S. C. Dance, Shauna H. Marr, Jacques Robert, Joseph E. Wedekind, Harold C. Smith and Andrea Bottaro

*J Immunol* 2006; 177:355-361; doi: 10.4049/jimmunol.177.1.355

http://www.jimmunol.org/content/177/1/355

References
This article cites 44 articles, 15 of which you can access for free at: http://www.jimmunol.org/content/177/1/355.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
Structural Phylogenetic Analysis of Activation-Induced Deaminase Function\textsuperscript{1,2}

H. Travis Ichikawa,\textsuperscript{3*} Mark P. Sowden,\textsuperscript{3*} Andrew T. Torelli,\textsuperscript{†} Jürgen Bachl,\textsuperscript{††} Pinwei Huang,\textsuperscript{†} Geoffrey S. C. Dance,\textsuperscript{4l} Shauna H. Marr,\textsuperscript{‡} Jacques Robert,\textsuperscript{‡} Joseph E. Wedekind,\textsuperscript{‡} Harold C. Smith,\textsuperscript{5†‡} and Andrea Bottaro\textsuperscript{5*‡§}

In mammals, activation-induced deaminase (AID) initiates somatic hypermutation (SHM) and class switch recombination (CSR) of Ig genes. SHM and CSR activities require separate regions within AID. A chromosome region maintenance 1 (CRM1)-dependent nuclear export signal (NES) at the AID C terminus is necessary for CSR, and has been suggested to associate with CSR-specific cofactors. CSR appeared late in AID evolution, during the emergence of land vertebrates from bony fish, which only display SHM. Here, we show that AID from African clawed frog (Xenopus laevis), but not pufferfish (Takifugu rubripes), can induce CSR in AID-deficient mouse B cells, although both are catalytically active in bacteria and mammalian cell systems, albeit at decreased level. Like mammalian AID, Takifugu AID is actively exported from the cell nucleus by CRM1, and the Takifugu NES can substitute for the equivalent region in human AID, indicating that all the CSR-essential NES motif functions evolutionarily predated CSR activity. We also show that fusion of the Takifugu AID catalytic domain to the entire human noncatalytic domain restores activity in mammalian cells, suggesting that AID features mapping within the noncatalytic domain, but outside the NES, influence its function. \textit{The Journal of Immunology, 2006, 177: 355–361.}

Immunoglobulin gene somatic hypermutation (SHM)\textsuperscript{6} and class switch recombination (CSR) require the activity of the B-cell-specific activation-induced deaminase (AID) enzyme (1, 2), a cytidine deaminase of the apolipoprotein B mRNA-editing catalytic polypeptide 1 (APOBEC-1)-related protein family (3, 4). AID also mediates gene conversion in chicken B cells (5, 6), with a mechanism that has been shown to be dependent on SHM induction (7). The leading model for AID function posits that cytidine (C) to uridine (U) transitions are directly introduced by AID into Ig variable gene segments and switch (S) regions DNA. Subsequently, the incorporated U residues at these loci are differentially processed through largely independent repair pathways, resulting in SHM and CSR (8, 9). A nonmutually exclusive hypothesis is that AID may act much like the prototypical member of the family, the RNA-editing enzyme APOBEC-1, to generate modified mRNAs encoding for one or more protein variants (e.g., a specific endonuclease), which would then directly mediate the initiation of CSR and/or SHM processes (10).

Unlike SHM and CSR, AID-dependent DNA deaminase activity in experimental systems shows limited sequence and locus specificity, although some site preference is detectable (8, 11–13). Specific sequence requirements for functional S regions are also not stringent (14–16). Based on analysis of AID mutants (17–19), and by analogy with other members of the APOBEC-1-related family of enzymes (4), it has been hypothesized that AID is directed to its natural substrates through its interaction with specific targeting factors. In particular, CSR activity has been linked to a C-terminal nuclear export signal motif (NES) in the noncatalytic domain of AID (17, 18), which mediates chromosome region maintenance 1 (CRM1)-dependent nuclear export (20–22). AID SHM activity appears to be dependent instead on separate motifs mapping to the catalytic domain of the molecule (19).

Phylogenetic evidence indicates that SHM and CSR appeared separately during vertebrate evolution (23, 24). SHM is known to occur already in cartilaginous fish, both in Ig and in nurse shark Ag receptor genes, albeit with variable patterns and frequencies, and is likely present in all jawed vertebrates (25–28). Class switching has been found so far only in land vertebrates (tetrapods), starting with amphibians (23, 24, 29). Although mechanistic differences between vertebrate classes are observed in SHM and CSR, such as in the pattern of mutation and the sequence of S region substrates, the underlying mechanisms are thought to be largely conserved (23, 24). AID homologs have been recently identified by sequence prediction or direct cDNA cloning in several teleosts and in shark.
isolated from Xenopus

To shed light on the phylogenetic and functional properties of AID during vertebrate evolution, we cloned AID cDNAs from Takifugu and Xenopus, and compared their structural and functional features with those of mammalian AID.

Materials and Methods

Cloning/sequences

Two complete Xenopus laevis AID (XaAID) cDNAs were cloned from two independent spleen cDNA libraries, using a PCR approach based on homologous sequences present in Assembly V2.0 of the Xenopus tropicalis genome (http://genome.igp-psf.org) (positions 6975–7397 of scaffold 44401) and the Xenopus EST project (www.genome.wustl.edu) database.

Takifugu rubripes AID (TrAID) cDNA was obtained by PCR from a T. rubripes spleen cDNA library (a gift from G. Elgar of the Fugu Genomics Group, London, U.K.) using DNA sequence data from Build 2c of the Takifugu genome (wwwensembl.org/Fugu_rubripes) for primer design.

The Takifugu-human AID fusion proteins were created by overlapping PCR from partial TrAID (aa 1–145) and Homo sapiens AID (HsAID) (aa 139–198) PCR fragments such that the flanking region of TrAID was retained.

All cDNAs were subcloned via pGEM7zf (Promega) and sequenced before subcloning into the appropriate vectors for functional analysis.

Analysis of AID expression in Xenopus

Outbred Xenopus were obtained from our breeding colony at the University of Rochester. For immunization, Escherichia coli (XLI-blue; Stratagene) were cultured overnight, boiled 30 min, spun, and resuspended in 0.1 vol of Xenopus cell culture medium (33). Two- to 3-year-old Xenopus were injected i.p. with 200 µl of this preparation. Total or cytoplasmic RNA was isolated from Xenopus tissue and 10–20 µg of RNA were analyzed by Northern blotting with a complete XAID cDNA as probe, or used for reverse transcription. For RT-PCR, samples were amplified using AID primers 5′-TCACGACCCCCCATAGAACCAC-3′, 5′-GGATTAGG AGACTTGGCCTCAAG-3′, or β2-microglobulin primers 5′-CCCTTGT GTGTTAAGTGTGTC-3′, reverse 5′-GCAACACACAACCTACGAA AAGGAC3′ for 40 cycles at 58°C annealing temperature.

AID-deficient mouse B cell line

All animal work was performed humanely according to federal and local guidelines and approved by the University Committee on Animal Research. DNA mutator activity by AID orthologs was assayed in BW310 strain E. coli using cDNAs cloned into plTrC99A vectors. Transformant clones were picked, grown in liquid culture, and induced with isopropyl β-D-thiogalactoside overnight before plating onto rifampicin selection or nonselective plates, essentially as described (8, 12).

Results

Cloning of T. rubripes and X. laevis AID and generation of fusion proteins

Takifugu and Xenopus AID cDNAs, were cloned from cDNA libraries based on genomic database information from the fugu and X. tropicalis genome projects. Compared with HsAID (37), XaAID contains two potential additional amino acid residues at the N terminus and an additional glycine at position 70 (human equivalent), for a total of 201 residues, a predicted molecular mass of 24 kDa, and an identity of 67.5% between the two species (Fig. 1A). XaAID mRNA expression in adult frogs is detectable mainly in spleen and was up-regulated following immunization with heat-killed E. coli (not shown), supporting a role for AID in humoral immune responses in amphibians analogous to that in mammals.

The cloned T. rubripes AID (TrAID) cDNA sequence encodes a 204-aa protein, of 24 kDa predicted mass, and 53% identical with HsAID, with a notable 8-residue insertion at position 75 (human equivalent) (Fig. 1A). A similar insertion was already noted in catfish and zebrafish, but not dogfish cDNA sequences (30–32). Based on our recently proposed three-dimensional model of AID, which predicts a tripartite structure comprising an N-terminal catalytic domain (CD), an active site flap, and a C-terminal noncatalytic domain (NCD) (38), the additional segment observed in TrAID would be located on the exterior of the molecule in a surface loop that connects the second helix region (α2) to the third β sheet (β3), distant from the active site, and therefore predicted not to affect catalytic activity. The first methionine in the cloned TrAID cDNA corresponds to residue 4 in mammalian AID, in agreement with the other Takifugu AID cDNA (32), but unlike the sequences predicted based on genome databases (30, 31). Regardless of whether the discrepancy is due to the presence of partial or alternative cDNA isoforms in Takifugu or to errors in sequence predictions, the observation that the TrAID catalytic domain is functional in mammalian cells shows that this specific difference does not significantly affect catalytic activity (see below).

To map potential salient structural-functional motifs onto the AID molecules, we also generated reciprocal fusion cDNAs in which the human and Takifugu CDs and flap regions were fused to the NCDs of the other species (TrCD/HsNCD and HsCD/TrNCD proteins) as well as a set in which 16-aa C-terminal amino acid segments, containing NES elements, were swapped (TrAID/HsNES and HsAID/TrNES) (Fig. 1B).

Mutator activity of TrAID and XaAID in bacteria

We evaluated whether the cloned cDNAs encode functional cytidine deaminase enzymes using a bacterial mutator assay, in which AID-mediated DNA deamination at cytidine residues results in increased frequency of rifampicin (Rif)-resistant bacterial colonies bearing mutations at the gene coding for the Rif target, RNA polymerase β (8, 12). XaAID expression in this system resulted in a...
small but significant increase in the generation of Rif-resistant mutants (~3-fold) compared with bacteria transformed by vector alone, although significantly lower than HsAID-expressing transformants (Fig. 2, A and B), despite comparable protein expression levels (not shown). TrAID was also capable of inducing a mutator phenotype in bacterial cells, ~4-fold higher than background (Fig. 2, A and B), which was further increased, as reported by Conticello et al. (32) by incubation of bacterial cultures at room temperature (not shown).

Analysis of the properties of the fusion proteins revealed that changes in the NCD can have striking effects on AID mutator activity. In particular, highly increased activity was observed for TrCD/HsNCD fusions, and to a lesser extent for TrAID/HsNES compared with wild-type TrAID (Fig. 2, A and B). Concomitant, but not as prominent decreases were observed for HsCD/TrNCD and HsAID/TrNES fusions compared with HsAID (Fig. 2, A and B). Together with the observation that mammalian AID truncation mutants missing the C-terminal region also have increased mutator activity in bacteria (17, 18), these data support the prediction from our proposed AID model that specific elements within the NCD, including but not limited to the C-terminal 16-aa region, may play a specific role in regulating substrate access to the AID catalytic site (38).

**CSR activity of AID orthologs and fusion proteins in mammalian cells**

To test for CSR activity of the various AID orthologs and fusions, cDNAs were expressed in LPS and IL-4-stimulated B cells from AID-deficient mice (1) by transfection with pMIG retroviral vectors (34), which express bicistronic transcripts also encoding

![Diagram showing the N-terminal Catalytic Domain (CD) and C-terminal NonCatalytic Domain (NCD) of AID orthologs and fusion proteins.](http://www.jimmunol.org/)

**FIGURE 2.** Functional analysis of AID orthologs and their fusions in bacterial DNA mutator assays. A, Expression of HsAID, TrAID, and XlAID, as well as HsAID/TrAID fusions promotes E. coli rifampicin resistance. Numbers of rifampicin-resistant (RifR) colonies in 10<sup>9</sup> total bacterial colonies are shown. Each dot represents an independent experiment. B, Expression of HsAID, TrAID, and reciprocal NES swap fusions similarly promotes rifampicin resistance in E. coli.
headings, section contents, and figures are not provided for the entire document. However, the text is clearly readable and makes sense as it stands. This text discusses the complementation of AID orthologs and their fusion proteins in B cells transduced with an AID-expressing retrovirus. The results suggest that AID activity is necessary for CSR in mammalian cells, and that the human NES motif can functionally replace its orthologous counterparts in other species. The text also mentions the use of bacterial mutator assays to assess AID activity, and the complementation of AID activity across different species, including human, mouse, and teleosts.

The text begins with a discussion of the experimental setup, where B cells were transduced with an AID-expressing retrovirus and analyzed for expression of IgG1 and EGFP. The results show that AID activity is limiting for CSR, and that the human NES motif can functionally replace its orthologous counterparts in other species. The text also mentions the use of bacterial mutator assays to assess AID activity, and the complementation of AID activity across different species, including human, mouse, and teleosts.

The text also discusses the activity of AID orthologs in mammalian cells, and the inability of TrAID to initiate CSR in mammalian cells could in principle be due to a lack of catalytic activity in mammalian cells, or to more specific differences in target recognition or regulation. The text concludes that all proteins are catalytically active in SHM, and that the correlation between the activities of each protein in SHM and CSR assays, results in bacterial mutator assays do not always match the activity of CSR in AID-deficient B cells (Fig. 3B, Table I), suggesting that elements critical to confer CSR activity to TrAID map to the mammalian C-terminal NCD. As CSR activity has been shown to require an active NES motif at the 16-aa C terminus of AID (17–19), we therefore analyzed the activity of reciprocal NES fusion proteins. The results showed that HsAID bearing the TrAID NES motif still retained significant CSR activity (~60% of wild-type) (Fig. 3C, Table I). In contrast, the reciprocal TrAID/HsNES fusion, with the human NES, could not induce significant CSR activity, because modest CSR activity was detectable in only one of three experiments (Fig. 3C, Table I). Taken together, these data suggest that elements critical for AID activity in this system map to the HsAID NCD upstream of the NES, but are absent in the equivalent region of TrAID.

Because an intact NES region is necessary for CSR induction by mammalian AID (17, 18) and, as shown above, the Takifugu AID NES can functionally replace its human counterpart in CSR, we considered whether it also shares with the mammalian NES the ability to mediate AID export from the nucleus via a CRM1-dependent mechanism (20–22). The NES region of AID from mammals, chicken, Xenopus, Takifugu and other teleosts (catfish and zebrafish) shows significant evolutionary conservation, particularly of the hydrophobic residues critical for nuclear export function (Fig. 4A). To test whether the Takifugu NES is indeed functional, we analyzed the activity of TrAID to be actively exported from the nucleus, as already shown for mammalian AID (20, 21). Like HsAID-EGFP fusion proteins, TrAID-EGFP displayed a predominantly cytoplasmic localization in transient NIH 3T3 cells, but accumulated in the nucleus upon treatment with the CRM1 inhibitor leptomycin B (Fig. 4B). We conclude that TrAID is actively exported from the nucleus via a CRM1-dependent mechanism.

**SHM activity of AID orthologs in mammalian cells**

The inability of TrAID to initiate CSR in mammalian cells could in principle be due to a lack of catalytic activity in mammalian cells, or to more specific differences in target recognition or regulation. We compared the ability of AID orthologs and fusions to induce SHM using a GFP gene stop-codon reversion assay in 293 cells (36). As shown in Fig. 5, all AID orthologs and fusions can induce detectable increases in GFP-stop reversion rates in this system, indicating that all proteins are catalytically active in SHM. Fig. 6 shows a summary of our observations with the various AID isoforms and fusion proteins. Although there is good correlation between the activities of each protein in SHM and CSR assays, results in bacterial mutator assays do not always match levels of CSR in AID-deficient B cells (Fig. 3B, Table I), suggesting that elements critical to confer CSR activity to TrAID map to the mammalian C-terminal NCD. As CSR activity has been shown to require an active NES motif at the 16-aa C terminus of AID (17–19), we therefore analyzed the activity of reciprocal NES fusion proteins. The results showed that HsAID bearing the TrAID NES motif still retained significant CSR activity (~60% of wild-type) (Fig. 3C, Table I). In contrast, the reciprocal TrAID/HsNES fusion, with the human NES, could not induce significant CSR activity, because modest CSR activity was detectable in only one of three experiments (Fig. 3C, Table I). Taken together, these data suggest that elements critical for AID activity in this system map to the HsAID NCD upstream of the NES, but are absent in the equivalent region of TrAID.

Because an intact NES region is necessary for CSR induction by mammalian AID (17, 18) and, as shown above, the Takifugu AID NES can functionally replace its human counterpart in CSR, we considered whether it also shares with the mammalian NES the ability to mediate AID export from the nucleus via a CRM1-dependent mechanism (20–22). The NES region of AID from mammals, chicken, Xenopus, Takifugu and other teleosts (catfish and zebrafish) shows significant evolutionary conservation, particularly of the hydrophobic residues critical for nuclear export function (Fig. 4A). To test whether the Takifugu NES is indeed functional, we analyzed the activity of TrAID to be actively exported from the nucleus, as already shown for mammalian AID (20, 21). Like HsAID-EGFP fusion proteins, TrAID-EGFP displayed a predominantly cytoplasmic localization in transient NIH 3T3 cells, but accumulated in the nucleus upon treatment with the CRM1 inhibitor leptomycin B (Fig. 4B). We conclude that TrAID is actively exported from the nucleus via a CRM1-dependent mechanism.

**Table I. CSR complementation in AID−/− mouse B cell by AID orthologs**

<table>
<thead>
<tr>
<th>Virus</th>
<th>IgG1+ Cells (% GFP+)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expt. 1</td>
</tr>
<tr>
<td>Vector</td>
<td>0.9</td>
</tr>
<tr>
<td>HsAID</td>
<td>37.4</td>
</tr>
<tr>
<td>XIaID</td>
<td>20</td>
</tr>
<tr>
<td>TrAID</td>
<td>0.5</td>
</tr>
<tr>
<td>TrAID/HsNCD</td>
<td>—</td>
</tr>
<tr>
<td>HsAID/TcNCD</td>
<td>—</td>
</tr>
<tr>
<td>TrAID/HsNES</td>
<td>—</td>
</tr>
<tr>
<td>HsAID/TcNES</td>
<td>—</td>
</tr>
</tbody>
</table>

b Frequency of IgG1+ cells among live-gated, B220+GFP+ day 5 lymphoblasts. — Not done.

A mouse AID-expressing retrovirus was used as positive control in this experiment.
with the activity in mammalian cells, clearly suggesting differential regulatory and functional constraints in prokaryotic vs eukaryotic systems.

**Discussion**

The evolutionary history of the adaptive immune system is characterized by an extraordinary plasticity in the addition and modification of genetic elements and functional modules, by mechanisms such as gene duplication, divergence, and co-option (9, 23, 24, 39, 40). AID well exemplifies this trend in its ability to mediate three mechanistically and phylogenetically distinct processes involved in the functional diversification of Igs: SHM in all jawed vertebrates, CSR in tetrapods, and gene conversion (GC) in (at least) birds (5, 6). The common mechanistic link between these three processes resides in AID’s activity as a cytidine deaminase (9, 10).

AID’s phylogenetic history must have involved several critical functional shifts. One is at its emergence as an enzyme for Ab variable region diversification, which must have occurred basally in jawed vertebrates, CSR in tetrapods, and gene conversion (GC) in (at least) birds (5, 6). The common mechanistic link between these three processes resides in AID’s activity as a cytidine deaminase (9, 10).

AID’s phylogenetic history must have involved several critical functional shifts. One is at its emergence as an enzyme for Ab variable region diversification, which must have occurred basally in jawed vertebrates, CSR in tetrapods, and gene conversion (GC) in (at least) birds (5, 6). The common mechanistic link between these three processes resides in AID’s activity as a cytidine deaminase (9, 10).

As this paper was in preparation, two groups have reported that zebrafish (Danio rerio) AID is capable of mediating normal CSR and SHM in mammalian cells (42, 43), whereas AID from both Takifugu (42) and catfish (Ictalurus punctatus) (43) have limited CSR and SHM activity. These data strongly argue that AID evolved the ability to mediate CSR before emergence of the class

**FIGURE 4.** Nuclear export activity of the TrAID NES. A, Alignment of the C-terminal amino acid regions of various AID orthologs highlights conservation of NES-critical motifs. B, CRM1-dependent nuclear export of TrAID. Localization of HsAID-EGFP (top panels) and TrAID-EGFP (bottom panels) in NIH 3T3 cells either untreated (left) or treated with the CRM1 inhibitor leptomycin B (LMB; right). Nuclei were stained with propidium iodide (PI). Two representative fields are shown for each experimental set. Arrows indicate AID-EGFP-expressing transfectants in fields containing multiple cells.

**Table II. Rate-limiting levels of AID-expressing retrovirus in transduced AID−/− mouse B cells**

<table>
<thead>
<tr>
<th>Virus</th>
<th>GFP Expression</th>
<th>Expt. 1</th>
<th>Expt. 2</th>
<th>Expt. 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GFP m.f.i.</td>
<td>IgG1+ (%)</td>
<td>GFP m.f.i.</td>
<td>IgG1+ (%)</td>
</tr>
<tr>
<td>HsAID</td>
<td>High</td>
<td>500</td>
<td>50</td>
<td>855</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>117</td>
<td>43</td>
<td>213</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>26</td>
<td>8.7</td>
<td>42</td>
</tr>
<tr>
<td>TrAID</td>
<td>High</td>
<td>515</td>
<td>0.4</td>
<td>868</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>122</td>
<td>0.4</td>
<td>212</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>29</td>
<td>0.4</td>
<td>42</td>
</tr>
</tbody>
</table>

*The frequency of IgG1+ cells was measured in live-gated, B220+ day 5 transduced lymphoblasts gated according to high, medium, or low GFP mean fluorescence intensity (m.f.i.).

on whether the known GC mechanisms in birds and mammals are homologous. These new functions could have been acquired by changes involving the AID protein itself, its ability to interact with regulatory cofactors, and/or by modifications in pathways downstream of AID activity (e.g., leading to differential repair of AID-induced DNA lesions). The latter may be the case for GC in birds, because recent evidence indicates that the mechanism is dependent on the differential resolution of AID-induced mutations in the primary Vγ segment in the presence or absence of GC donor dV sequences (7). The same has been proposed for CSR, based on the observation that S region mutations accompany class switching in vitro (41). In contrast, the finding that separate regions within AID are required for CSR and SHM, possibly via association with distinct cofactors, supports the hypothesis that specific changes in the AID protein itself may have been involved in CSR emergence (17–19).

As this paper was in preparation, two groups have reported that zebrafish (Danio rerio) AID is capable of mediating normal CSR and SHM in mammalian cells (42, 43), whereas AID from both Takifugu (42) and catfish (Ictalurus punctatus) (43) have limited CSR and SHM activity. These data strongly argue that AID evolved the ability to mediate CSR before emergence of the class

**FIGURE 4.** Nuclear export activity of the TrAID NES. A, Alignment of the C-terminal amino acid regions of various AID orthologs highlights conservation of NES-critical motifs. B, CRM1-dependent nuclear export of TrAID. Localization of HsAID-EGFP (top panels) and TrAID-EGFP (bottom panels) in NIH 3T3 cells either untreated (left) or treated with the CRM1 inhibitor leptomycin B (LMB; right). Nuclei were stained with propidium iodide (PI). Two representative fields are shown for each experimental set. Arrows indicate AID-EGFP-expressing transfectants in fields containing multiple cells.
switch mechanism, although the molecular basis for the observed differences between fish species was not addressed. Together with data indicating that *Xenopus* S regions are recognized as switch targets by mammalian AID despite significant sequence divergence (16), the observation of zebrafish AID CSR activity and our finding that *Xa* AID is capable of replacing mammalian AID to initiate CSR in activated B lymphocytes indicate a strict conservation of the entire CSR mechanism over many hundreds of millions of years and across divergent vertebrate lineages.

The inability of *Takifugu* AID to complement AID-deficient mouse B cells in our experiments is also in substantial agreement with the recent reports, although in our hands the deficiency is more pronounced than what was observed by Barreto et al. (42) with Takifugu, and by Wakae et al. (43) with catfish proteins. Although we cannot presently pinpoint the cause of this difference, we note that our retroviral AID expression system is not at functional saturation, as shown by the linear correlation between GFP expression and switch efficiency in transduced B cells (Table II), and may therefore be more sensitive to variations of activity. Together, however, the data presented here and from other labs (41, 42) paint a consistent picture of significant variation in the ability of AID from various fish species to complement mammalian AID activity in SHM and CSR. Simple variation in expression levels and/or stability are unlikely to account for the functional differences, because comparable expression levels are found in mammalian cells for *Takifugu* and zebrafish AID (42), mouse and catfish AID (43), and human and *Takifugu* AID (our data; not shown).

From an evolutionary standpoint, zebrafish and catfish are more closely related to each other than to *Takifugu* (44), with tetrapods, which include *Xenopus* and mammals, as an outgroup. This suggests two alternative, equally parsimonious evolutionary scenarios, i.e., that zebrafish and tetrapods acquired their activity independently, or that *Takifugu* and catfish independently lost theirs. Regardless of whether the activity differences between AID orthologs are due to independent gain or loss of function, however, convergent evolution of this kind would seem more plausible based on quantitative/regulatory differences, e.g., ability to associate with cofactors, as opposed to structural changes affecting AID catalytic properties such as substrate specificity, but more detailed studies will be required to address this issue.

Our finding that replacement of the *Tr* AID noncatalytic domain with the human counterpart restores activity in mammalian cells rules out that the difference between *Takifugu* and mammalian AID involves critical differences in the AID catalytic site, or other elements in the catalytic domain, such as the peculiar teleost-specific 8-residue insertion, or the absence of some N-terminal residues in *Tr* AID compared with mammalian AID. The primary candidate for any novel element in the AID NCD that could have been evolutionarily acquired during CSR emergence is the C-terminal NES motif (17, 18). Although the only confirmed NES function is CRM1-mediated nuclear export (20, 21), this element has been proposed to also mediate interaction with CSR-specific cofactors, acting either at the targeting stage, or after the AID-induced C-to-U DNA deamination (17, 18). Importantly, our results clearly show that both CRM1-mediated nuclear export and any additional putative CSR-specific function of the AID NES must have actually existed before the onset of CSR.

This finding implies that AID nuclear export must also play a selectable CSR-unrelated, ancestral function in fish. Because it has been shown that nuclear export may limit AID’s activity as a generalized mutator (21), it is possible for instance that the AID NES originally evolved as a control mechanism for potential genotoxicity. The role of the NES in switching may represent therefore a form of molecular preadaption, later exploited for CSR evolution. How could the NES evolve for one function, and later be co-opted and become necessary for CSR? The first possibility is that CRM1 itself, or a factor with the same binding specificity for the NES motif as CRM1, acts as a CSR cofactor. Alternatively, the NES role in CSR may be indirect, that is, CSR evolution may have required the act of nuclear export per se. Perhaps nuclear export allowed AID to target a new substrate by homing to a specific compartment, to undergo posttranslational modifications (e.g., phosphorylation by a cytoplasmic kinase), or to associate with cofactors in the cytoplasm. The latter would imply the existence of at least one yet-uncharacterized CSR-specific element within AID.

The ability of the *Hs* AID NCD to confer high activity to the *Tr* AID CD in mammalian and bacterial cells is also of interest. Although it is possible that the effect is entirely mediated by structural/folding properties of the protein, it would seem rather peculiar for a *trans*-specific fusion protein to be structurally more favorable than either wild-type forms. A more interesting possibility is that the effect relates to specific regulatory function(s) of the NCD over AID activity, either by restricting substrate access, or by regulating catalytic specificity. Thus, any more “relaxed” protein configuration may increase overall catalytic activity, in agreement with observations on the catalytic effect of both C-terminal truncations (17, 18) and NES/NCD swaps (this paper) and with the proposed AID model (38). This would raise the intriguing possibility that AID was evolutionarily under strong selection for catalytic suboptimality, again, as discussed above, because of its potential mutator role.

Further studies, combining more extensive functional phylogenetic analysis and structural modeling will be required to highlight
Disclosures

The authors have no financial conflict of interest.

References


