Specific Expression of Activation-induced Cytidine Deaminase (AID), a Novel Member of the RNA-editing Deaminase Family in Germinal Center B Cells*

(Received for publication, January 12, 1999, and in revised form, March 19, 1999)

Masamichi Muramatsu‡, V. S. Sankaranand‡, Shrikant Anant‡, Manabu Sugaï, Kazuo Kinoshita‡, Nicholas O. Davidson§, and Tasuku Honjo‡

From the ‡Department of Medical Chemistry, Faculty of Medicine and §Center for Molecular Biology and Genetics, Kyoto University, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan and ‡Department of Internal Medicine, Division of Gastroenterology, Washington University School of Medicine, St. Louis, Missouri 63110

We have identified a novel gene referred to as activation-induced deaminase (AID) by subtraction of cDNAs derived from switch-induced and uninduced murine B lymphoma CH12F3-2 cells, more than 80% of which switch exclusively to IgA upon stimulation. The amino acid sequence encoded by AID cDNA is homologous to that of apolipoprotein B (apoB) mRNA-editing enzyme, catalytic polypeptide 1 (APOBEC-1), a type of cytidine deaminase that constitutes a catalytic subunit for the apoB mRNA-editing complex. In vitro experiments using a glutathione S-transferase AID fusion protein revealed significant cytidine deaminase activity that is blocked by tetrahydrouridine and by zinc chelation. However, AID alone did neither demonstrate activity in C to U editing of apoB mRNA nor bind to AU-rich RNA targets. AID mRNA expression is induced in splenic B cells that were activated in vitro or by immunizations with sheep red blood cells. In situ hybridization of immunized spleen sections revealed the restricted expression of AID mRNA in developing germinal centers in which modulation of immunoglobulin gene information through somatic hypermutation and class switch recombination takes place. Taken together, these findings suggest that AID is a new member of the RNA-editing deaminase family and may play a role in genetic events in the germinal center B cell.

The germinal center (GC) constitutes a highly specialized microenvironment required for the final maturation step of naive B cells toward antigen-specific memory cells or long-lived plasma cells (1, 2). Two prominent alterations of immunoglobulin gene information are known to occur in this microenvironment (3–5). First, accumulation of massive point mutations in the variable region exon, a process refer to as somatic hypermutation (6), gives rise to affinity maturation of antibody in association with selection of B cells expressing high affinity immunoglobulins on their surface (7, 8). Second, class switch recombination (CSR) replaces the exons encoding the heavy chain constant region (9), which determines effector functions of the antibody including complement fixation. These two alterations of immunoglobulin gene information are critical for accounting for an effective humoral response to harmful microbes. The molecular mechanisms for these genetic events remain to be elucidated despite intensive study.

To dissect the molecular mechanism of class switching, we have isolated a murine B lymphoma clone CH12F3-2 in which CSR from IgM to IgA begins to occur within a few hours after stimulation with IL-4, TGF-β, and CD40L, giving rise to more than 80% IgA+ cells (10–12). Using CH12F3-2 cells, we have shown that CSR breakpoints are distributed not only within typical repetitive sequences (designated S region) but also in its flanking regions (11). However, the break points were rarely found in the I and C exons, which are separated by the S region, i.e. the intron of germ-line transcripts. Because accumulating evidence indicates that transcription from I to C exon and splicing of the transcripts are essential to CSR (13, 14), it is possible that the transcripts are involved directly or indirectly in CSR. We have thus proposed that a complex structure of DNA and RNA, but not S region sequence per se, is recognized for initiation of CSR (11). This idea gained further support with the finding that CSR can take place efficiently even when the Sα region was replaced by the Sε or Sγ1 region in minichromosome constructs introduced in CH12F3-2 cells upon cytokine stimulation (10).

Another type of genetic regulation, RNA-editing, is widely used as a means to create new functional genes from the restricted genome in plants and protozoa (15, 16). An increasing number of mammalian mRNAs are also known to be edited, including apolipoprotein B (apoB) mRNA, glutamate receptor mRNA, Wilms tumor-1 mRNA, α-galactosidase mRNA, neurofibromatosis type-1 mRNA, and tRNA<sup>α</sup> (17). Although most of their molecular mechanisms are yet to be elucidated, that of apoB mRNA editing by APOBEC-1(18, 19) is extensively documented. ApoB mRNA editing involves a site-specific C to U deamination of the first base of a CAA codon, encoding glutamine at residue 2153 in apoB100, and produces a UAA frame stop codon in apoB<sub>48</sub> mRNA (20). ApoB100 and apoB<sub>48</sub> are translation products of the unedited and edited apoB mRNAs, respectively, and these proteins have completely different physiological functions (21). APOBEC-1 requires an aux-
The report describes the isolation and characterization of AID, a gene that is specifically expressed in GC B cells. The restriction enzyme-panel PCR amplification identified AID as a gene enriched for inducible genes upon CSR according to the manufacturer’s instructions. The CH12F3-2 cells were cultured and stimulated as described before (12, 25). Poly(A) RNAs for a tester and a driver cDNA were isolated from stimulated and unstimulated CH12F3-2 cells, respectively. HaeIII-digested 4X174 phage DNA was added to the tester DNA to monitor the efficiency of subtraction. The subtracted cDNA was cloned into T-vector (Promega), and the plasmid library was generated. Two thousand colonies from the subtracted library were screened by differential hybridization as described previously (25). The tester cDNA probe used in this study were derived from CH12F3-2 cells stimulated for 5 and 12 h. To discriminate clones containing HaeIII-digested 4X174 phage DNA, replica filters were hybridized with HaeIII-digested 4X174 DNA. After subtraction, the final cDNA pool was 100-fold enriched. Plasmids from 115 positive colonies that showed stronger signals with the cDNA probe derived from stimulated cells were sequenced, and their inserts were subjected to Northern blot hybridization to demonstrate up-regulated expression upon CSR in CH12F3-2 cells. Twenty-three cDNA fragments were found to be up-regulated by stimulation with IL-4, TGF-β, and CD40L, including 8 ABCD-1/MDC, 2 interferon-γ receptor, 1 I-a (major histocompatibility complex class II), 3 AID, 7 15B11, 1 B89, and 1 16A9. Full-length cDNA clones encoding AID gene was obtained by cDNA library screening (25).

Production of GST Fusion Protein and Cytidine Deaminase Activity—AID cDNA was subcloned into pGEX4T1 vector (Amer sham Pharmacia Biotech) after PCR amplification with AID-138 and AID-161 primers and sequenced on both strands. AID cDNA was expressed as a glutathione S-transferase (GST) fusion protein in Escherichia coli. The fusion protein was purified essentially as described previously (22) with the exception that the GST fusion protein was purified by two rounds of glutathione-agarose chromatography. The purified GST fusion protein was confirmed by Western blot analyses as described previously (26) using anti-AID antisera, which were prepared by immunizing rabbits with a multiple-antigen peptide (27) containing residues 185-198 (EVDLRLDRAFRMLGF). Cytidine deaminase activity was determined as described previously (22).

Mouse Immunization and Fractionation of Splenic Cells—BALB/c mice (6–12 weeks of age) were purchased from Japan SLC Inc. (Shizuoka, Japan). Purified splenic B cells were obtained and stimulated as described previously (25). In this study, dead cells were removed by Ficoll gradient after T-cell depletion and 50 µg/ml LPS from Salmonella typhosa (Sigma) were used to stimulate cells. Mice were intraperitoneally immunized with 1×10^6 sheep red blood cells (SRBC; Cosmo Bio). Nylon fiber (Wako Chemicals) was used to separate T cells from RBC-depleted splenic cells (28). The final T cell fraction contained more than 90% CD3+ cells and less than 3% B220+ cells. To enrich or deplete B cells, magnetic beads-conjugated anti-CD19 antibody (Miltenyi Biotec) was applied, and the mixture was passed over anolumn. The CD19+ cells depleted fraction contained less than 5% B220+ cells, whereas the CD19+ cell-enriched fraction contained more than 60% B220+ B cells.

In Situ Hybridization—An EcoRI/Xhol fragment of AID cDNA was subcloned into pBluescriptSK(+) (Stratagene). The plasmid was linearized with EcoRI or Xhol and used for RNA transcription with T3 or T7 RNA polymerase to generate antisense and sense probes, respectively, in the presence of digoxigenin-labeled rUTP (Roche Molecular Biochemicals). In situ hybridization procedures and detection of riboprobes were conducted as described previously (29). The slides were examined under a standard light transmission microscope.

Northern Acid Analyses—PCR amplification of looped-out circular DNA-containing S-regions were performed as described previously (11) using αF1 and μR3(first) primers for primary PCR and αF1 and μR3(second) primers for secondary PCR. Total RNA and poly(A)+ RNA were purified as described previously (25). Northern and Southern blot hybridizations were performed according to standard procedures. Amounts of AID mRNA were radioquantitated from Northern blot data after normalization with intensities of GAPDH bands and represented as relative folds of nonstimulated samples. RT-PCR was carried out as described previously (10). GAPDH cDNA was amplified using GF and GR primers. AID cDNA was amplified using AID-118 and AID-119 primers. Because an intron exists between AID-118 and AID-119 primers, the PCR product derived from contaminating genomic DNA was easily distinguished. Sequence comparison was performed with GENETYX MAC 9.0, and a data base search was done with BLAST (30) using GenBank® as DNA data bases and SwissProt as protein data base. Motif search analysis was performed on line at Prosite (31).

RESULTS

Construction and Screening of CH12F3-2 Cell cDNA Libraries Enriched for Gene Sequences Induced upon CSR—Within hours of stimulation of CH12F3-2 cells with IL-4, TGF-β and CD40L, it is possible to detect looped-out circular DNA containing S regions, which represent the earliest signs of CSR (11). Because loop-out of the circular DNA was inhibited with cycloheximide, we reasoned that up-regulation of proteins is required in a very early phase after stimulation for CSR to take place (Fig. 1, A and B). To isolate such early response genes with CSR induction, we performed PCR-based subtraction using cDNAs obtained from CH12F3-2 cells, shortly after and before stimulation. 115 clones that showed significantly stronger signals for the stimulated probe compared with the non-stimulated probe were sequenced. Seven independent sequences were obtained from the 115 clones, all of which were up-regulated upon stimulation in CH12F3-2 cells, as confirmed by Northern blot hybridization. The seven sequences contained four novel genes and three known genes, which include I-a (major histocompatibility complex class II), ABCD-1/MDC, and interferon γ receptor. I-a and ABCD-1/MDC are known to be up-regulated by IL-4 and CD40L in splenic B cells (32), indicating that the subtraction was effective. Among four novel genes, a clone encoding AID was selected for further analysis because its basal expression in nonstimulated CH12F3-2 cells was low, whereas drastic up-regulation (10 folds at 6 h) was observed upon stimulation (Fig. 1C).

Isolation and Structure of a Novel Gene, AID—To obtain a full-length AID cDNA clone, a 1020-base pair AID fragment was used as a probe for screening a cDNA library of stimulated CH12F3-2 cells. One clone (1.2 kb) revealed a single open reading frame and one polyadenylation site. Two major transcripts (2.4 and 1.2 kb), observed in Northern blot (Fig. 1C), appear to
correspond to the two species of cDNA utilizing the 3' and 5' polyadenylation sites, respectively.

The open reading frame of AID cDNA encodes a 198-residue protein with a predicted molecular mass of 24 kDa. Homology search against data bases of known protein sequences revealed significant homology (34% amino acid identity) with mouse APOBEC-1 (Fig. 2). A data base search revealed the presence of an active site for cytidine deaminases, which is conserved in the large cytosine deaminase family. The family can be classified into RNA-editing deaminases, cytidine deaminases, and CMP deaminases, based on substrate specificity and homology of the active-site sequence (20). As shown in Fig. 2B, the phylogenetic tree shows that the active site of AID is closer to the subgroup of RNA-editing deaminases rather than those of cytidine and CMP deaminases. The leucine-rich region located in the C terminus of APOBEC-1 has been proposed to be important for protein-protein interaction (24, 33). The AID amino acid sequence also contains the C-terminal leucine-rich region, in which four leucines of AID are conserved among rabbit, rat, mouse, and human APOBEC-1 (Fig. 2A).

Cytidine Deaminase Activity of the AID Protein—Recombinant APOBEC-1 is known to catalyze in vitro deamination of apoB mRNA (multimeric substrate) as well as cytidine (monomeric substrate) (20). Because the cytidine deaminase motif in AID is completely conserved, we investigated the cytidine deaminase activity of AID using a purified GST-AID fusion protein. As shown in Fig. 3A, the purified GST-AID fusion protein migrated to the expected position on SDS-polyacrylamide gel (49 kDa). The purified protein was confirmed by an anti-AID peptide antibody recognizing the C terminus of AID.

**Fig. 2.** Alignment of amino acid sequences of AID and mouse APOBEC-1 and the phylogenetic tree of the cytosine deaminase family. A, identical amino acids are marked by asterisks. Hyphens represent gaps in the sequence generated by the program to allow better alignment. The conserved active site of cytidine deaminases is indicated by an open box. Arrowheads represent phenylalanines and leucines, which are conserved among rat, mouse, rabbit, and human APOBEC-1. Underlines indicate the pseudoactive domain. B, The phylogenetic tree was derived by alignment of the active site of cytidine deaminases, using the UPGMA method. Accession numbers of amino acid sequences used here are as follows; human cytidine deaminase, L27943; mouse cytidine deaminase, AA388666; Bacillus subtilis cytidine deaminase, U18532; E. coli cytidine deaminase, X63144; rabbit APOBEC-1, U10695; human APOBEC-1, L25877; rat APOBEC-1, U10695; mouse APOBEC-1, U21951; T2/T4 phage CMP deaminase, J05172; human CMP deaminase, L12136; Saccharomyces cerevisiae CMP deaminase, U10397.

**Fig. 1.** Inhibition of CSR and AID induction by cycloheximide (CHX). CH12F3-2 cells were treated with or without stimulants (IL-4, TGF-β, and CD40L) in the presence or absence of 200 ng/ml CHX for 6 h. Genomic DNAs were extracted and subjected to PCR to amplify circular DNA containing Sα and Sα sequences or GAPDH genomic DNA. A, ethidium bromide (EtBr) staining of amplified DNA. Formation of switch circular DNA is inhibited by CHX treatment. M, molecular mass standards. B, Southern blot hybridization of the same gel in A with Sα probe. C, CHX abrogates the up-regulation of AID mRNA. CH12F3-2 cells were treated for 6 h with either medium, CHX (200 ng/ml), stimulants, or CHX plus stimulants. Northern blot hybridization of 10 μg of total RNA was carried out with AID and GAPDH cDNA probes. Induction fold of AID mRNA expression is indicated below.
protein. Minor bands seen below the major band for GST-AID in silver staining may be degradation products, some of which were detected by Western blot with the anti-AID antibody. Fig. 3B demonstrates that the cytidine deaminase activity of the AID fusion protein increases in a dose-dependent manner. \( K_m \) and \( V_{max} \) values were calculated to be 1.6 \( \times 10^{-5} \) M and 1.33 unit/mg, respectively. The deaminase activity of AID was inhibited by a cytidine deaminase-specific inhibitor, tetrahydouridine (THU) (Fig. 3C). The activity is also sensitive to a zinc chelator, 1,10-o-phenanthroline, with 91% inhibition at 20 mM but only 13% inhibition by its inactive isomer, 1,7-o-phenanthroline, indicating that, like APOBEC-1, AID is a zinc-dependent cytidine deaminase (Fig. 3D). Recombinant APOBEC-1 is known to bind AU-rich RNA (17, 20, 22, 23) and to mediate apoB RNA-editing in the presence of chicken extracts containing auxiliary factors. An RNA substrate of APOBEC-1 was challenged by AID to test its RNA-editing activity. Although AID is a functional cytidine deaminase with structural similarity to APOBEC-1, a gel retardation assay showed no binding capacity of AID to the AU-rich RNA, and the in vitro apoB RNA-editing assay (22) showed no C to U conversion (data not shown).

Specific Induction of AID mRNA Expression in Activated B Cells in Lymphoid Organs—Northern blot analysis of different mouse tissues revealed that AID mRNAs were abundant in mesenteric lymph nodes, scarce in spleen, and undetectable in other nonlymphoid tissues (Fig. 4A). Expression of AID mRNAs in various lymphoid organs was further investigated by RT-PCR (Fig. 4B). Transcripts for AID were observed in peripheral lymphoid organs such as Peyer’s patches and lymph nodes, with lower levels in bone marrow. AID mRNA was not detected in thymus.

The time course of AID mRNA induction in CH12F3-2 cells was further determined by Northern blot hybridization (Fig. 5A). Induction of AID mRNAs was detectable by 3 h after stimulation with IL-4, TGF-\( \beta \), and CD40L, reached a maximal level (13–14-fold induction) at 12 h, and declined gradually after 48 h. To determine which stimulant was responsible for induction of AID mRNAs, different combinations of reagents were tested for expression of AID mRNAs by Northern blot hybridization (Fig. 5B). Although a single stimulation with either IL-4, TGF-\( \beta \), or CD40L slightly induced AID expression, simultaneous addition of the three was most effective as previously shown for CSR induction from IgM to IgA in CH12F3 cells (10, 12). Up-regulation of AID mRNA appears to require de novo synthesis of proteins, because its induction was profoundly reduced by cycloheximide treatment (Fig. 1C).

We next investigated whether AID mRNAs are induced in splenic B cells by reagents shown to activate B cells and induce CSR. Fig. 6 demonstrates that AID mRNAs were strongly up-regulated in splenic B cells cultured in the presence of either LPS, LPS + IL-4, or LPS + TGF-\( \beta \). Stimulation of splenic B cells with CD40L resulted in AID mRNA induction, albeit to a lower level. We further investigated whether expression of AID mRNAs is up-regulated by antigen stimulation in vivo. An immunization with SRBC is known to initiate immunological responses, followed by clonal expansion and germinatal center formation, in which CSR and affinity maturation take place. Poly(A)+ RNA was extracted from spleens of mice immunized with SRBC and subjected to Northern blot hybridization (Fig. 7A). AID mRNA expression was barely detectable in spleens without immunization, and significant up-regulation (3–5-fold induction) was demonstrated 5 and 13 days after immunization. To determine which cells express AID mRNAs in immunized spleen, RT-PCR was performed using fractionated splenic cells. Fig. 7B shows that AID mRNAs were observed in the non-T and CD19+ cell fractions, indicating that SRBC immunization induced AID mRNA expression in CD19+ splenic B cells.

Specific Expression of AID mRNA in Germinal Centers—Because the timing of AID mRNA up-regulation in the spleen was roughly synchronous with the onset of GC formation after immunization (Fig. 7A), we next investigated the precise local-
indicated below. A, CH12F3-2 cells were treated by IL-4, TGF-β, and CD40L for indicated periods. B, specificity of stimulants were tested by stimulation of CH12F3-2 cells with indicated stimulants for 12 h.

Fig. 5. Induction of AID mRNAs in CH12F3-2 cells. Each lane contains 10 μg of total RNA from nonstimulated or stimulated CH12F3-2 cells. Northern blot filters were hybridized with the AID and GAPDH cDNA probes. Induction fold of AID mRNA expression was indicated below. A, CH12F3-2 cells were treated by IL-4, TGF-β, and CD40L for indicated periods. B, specificity of stimulants were tested by stimulation of CH12F3-2 cells with indicated stimulants for 12 h.

Fig. 6. Induction of AID mRNAs in splenic B cells stimulated in vitro. Purified murine splenic B cells were cultured with the indicated stimulants for 4 days. Total RNA (15 μg) from stimulated cells were electrophoresed, stained with ethidium bromide, and analyzed by Northern blot hybridization with an AID cDNA probe. The same filter was hybridized with a GAPDH cDNA probe.

Fig. 7. Induction of AID mRNAs in vivo. A, induction of AID mRNAs in murine splenic B cells following immunization with SRBC. Mice were intraperitoneally immunized with SRBC. At the indicated time, five spleens were resected and pooled to extract RNA, and 2 μg of poly(A)⁺ RNA was isolated and subjected to Northern blot hybridization. Induction fold of AID mRNA expression was indicated below. B, expression of AID mRNAs in fractionated splenic B cells obtained from mice immunized for 5 days. Spleens were resected 5 days after immunization with SRBC, and the splenic cells were fractionated by either MACS using anti-CD19 antibody (1st and 2nd lanes) or nylon-fiber separation (3rd and 4th lanes). RNA from each fraction was subjected to RT-PCR.

Among four novel genes thus isolated, we have characterized AID, which is specifically induced in GC B cells upon immunization. AID contains the active site for cytidine deaminase and catalyzes deamination of cytidine in vitro (Fig. 3). The inhibitory effect of THU and of zinc chelation on enzyme activity suggests that the deamination process may be similar to that of other cytosine deaminases, including APOBEC-1(19). Phylogenetic analysis revealed that AID is located closer to the RNA-editing deaminase rather than other cytosine deaminases (Fig. 2B), despite the fact that AID lacks RNA editing activity on an apoB RNA template. Mutagenesis studies indicate that Phe-66, Phe-87, His-61, Glu-63, and Cys-93 in mouse APOBEC-1 are essential for RNA binding (17, 20, 22, 23). These residues were noted to be conserved in the AID primary structure (Fig. 2A). However, as noted above, AID was not found to exhibit apoB RNA binding activity. X-ray crystallographic studies on the E. coli cytidine deaminase has revealed its three-dimensional structure, which has similarities with the predicted structure of APOBEC-1(34). The cytosine deaminase family can be divided into two groups in the view of the quaternary organization (35). E. coli cytidine deaminase and APOBEC-1 have a pseudoactive site domain in the C terminus, which is required to form a homodimer, whereas the other group, e.g. CMP deaminase, lacks such domain and forms monomers. AID contains a leucine-rich C terminus, which is shorter than but similar to the pseudoactive site domain of APOBEC-1 (Fig. 2A). AID belongs to the APOBEC-1/E. coli cytidine deaminase group. The predicted structural similarity of AID to APOBEC-1 implies that AID may be a novel RNA-editing deaminase induced in GC B cells, although it remains to be determined whether AID assembles into either a homodimer or associates into a hetero-heteromeric with the auxiliary factors used by APOBEC-1. Further resolution of this issue will require formal identification of the auxiliary factors themselves.

AID mRNA was induced in CH12F3-2 cells within 3 h after cytokine stimulation (Fig. 5). The appearance of AID mRNAs and the onset of CSR in CH12F3 cells thus coincide temporally. AID mRNA expression was induced in splenic B cells after in vitro treatment with LPS and cytokine, which activates naive B cells to initiate CSR (Fig. 6). AID mRNAs were also up-regulated in splenic B cells in vivo when mice were immunized with the T cell-dependent antigen, SRBC. Furthermore, ex vivo experiments (Fig. 7) and in situ hybridization (Fig. 8) studies

DISCUSSION

To dissect the molecular mechanism of CSR, we screened subtracted cDNA libraries from nonstimulated and stimulated CH12F3-2 cells on the assumption that induction of transacting factors, such as switch recombine, are required for CSR. In support of this, it was demonstrated that cycloheximide treatment inhibited the formation of looped-out circular DNA in stimulated CH12F3-2 cells (Fig. 1, A and B), indicating the requirement for de novo protein synthesis for CSR to take place.
revealed that AID mRNA was specifically detected in GC B cells, which are competent to perform somatic hypermutation and CSR. In addition, none of the cell lines examined that do not support CSR, including LyD9, BA/F3, 70Z/3, WEHI231, X63, WEHI-3, EL-4, 2B4, F2, P815, L929, NIH3T3, and ST2, were found to express AID mRNA (data not shown). Taken together, these data indicate that AID mRNA is a GC B cell-specific gene induced upon antigen stimulation. Highly restricted expression of the AID gene in GC B cells together with the concordant onset of AID induction and CSR suggest its role in GC function. Furthermore, the possibility of RNA-editing activity on other templates inspires us to speculate that AID may participate in regulatory steps unique to GC function such as somatic hypermutation and CSR. It is of note that the efficiency of CSR in CH12F3-2 cells is increased up to double when cells were stimulated in the presence of THU (data not shown). The physiological function of AID is currently being investigated by gene targeting and transgenic techniques.

Acknowledgments—We thank Drs. H. Ikuta S. Minoguchi, and K. Kuroda for discussion and suggestions. We also thank M. Yamamoto and T. Tabuchi for their technical assistance and Y. Takahashi and T. Tanaka for secretarial help.

REFERENCES

13. Hein, K., Lorenz, M., Siebenkotten, B., Petry, K., Christine, R., and Radbruch,