Studies on Functions of SRSF1-3 in Somatic Hypermutation Specific to the Immunoglobulin Gene

March, 2017

Yuka Kawaguchi

The Graduate School of

Natural Science and Technology

(Doctor Course)

OKAYAMA UNIVERSITY

Table of Contents

General Introduction

Generation of Antibody Repertoire and Antibody Affinity Maturation in the Immune Response

We have two biological defense mechanisms in the immune system to exclude pathogens that invade into our bodies. The first line of defense is the innate immunity that constitutively functions to immediately exclude pathogens by recognizing the molecular patterns on the pathogens. The second wave is the adaptive immunity that targets specific pathogens that break down innate immune barriers. The adaptive immunity in higher organisms, i.e., vertebrates, has systems that can recognize a wide variety of antigens to defend against many kinds of pathogens by generating a lot of antigen receptors, each of which can specifically recognize its target. B cells are involved in one of the systems.

B cells express B cell antigen receptors (BCR) on the cell surface and antibody molecules that are a secreted form of BCR. The antibody consists of the variable region that binds antigens and the constant region that is structurally conserved between antibody molecules (Fig. 1). Each immunoglobulin variable (IgV) region is capable of recognizing different antigen(s) because IgV is generated by recombination of multiple V segments, D segments, and J segments with the activity of the recombinase encoded in the recombination activating gene 1 (RAG1) and RAG2 (1). As a result of the V(D)J recombination, naïve B cell population can be composed of more than 10^{11} kinds of various antibody repertoire. Naïve B cells circulate through the blood stream, and when they encounter antigens that bind to their BCRs, they are activated and in addition receive activation and differentiation signals from dendritic cells and T cells to differentiate into plasma cells and memory B cells (2). In the course of the B cell response, the affinity and the effector functions of the antigen-specific antibodies produced by the activated B cells are increased. The mechanism is called antibody affinity maturation.

The affinity maturation depends on activation-induced cytidine deaminase (AID). AID damages DNA by converting dC into dU in DNA, and specifically targets the Ig gene locus (3).

When AID acts on the IgV region gene, somatic hypermutation (SHM) is induced. In some animals, such as birds and rabbits, DNA lesions by AID trigger gene conversion (GCV) through the homologous recombination repair mechanism, which templates pseudogenes existing upstream of the IgV region gene. On the other hand, when AID acts on the switch region upstream of the Ig constant region genes, recombination occurs between two switch regions and splices out the constant region gene between the switch regions, thereby recombining the IgV gene and the downstream constant region gene to be expressed (Fig. 2). Through this mechanism, the class of the antibody constant region changes, which is called class switch recombination (CSR). Taken together, AID contributes to the secondary diversification necessary for the antibody affinity maturation and the change of antibody effector functions (4) (5).

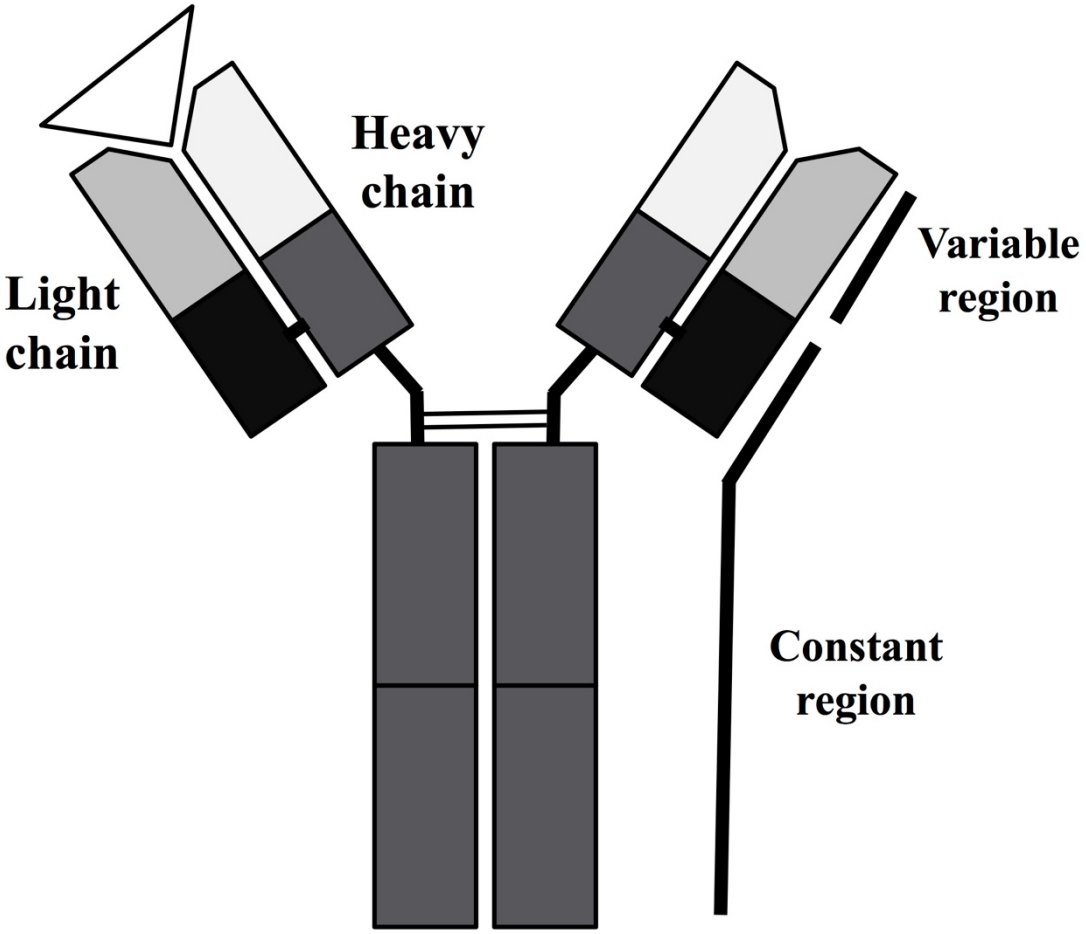


Figure 1. Antibody is composed of the variable region and constant region.

The variable (V) region is generated from combinations of gene segments to have a diversified antigen binding site for various antigens. There are five classes of antibodies, whose constant (C) regions are different: IgM, IgD, IgG, IgA, and IgE.

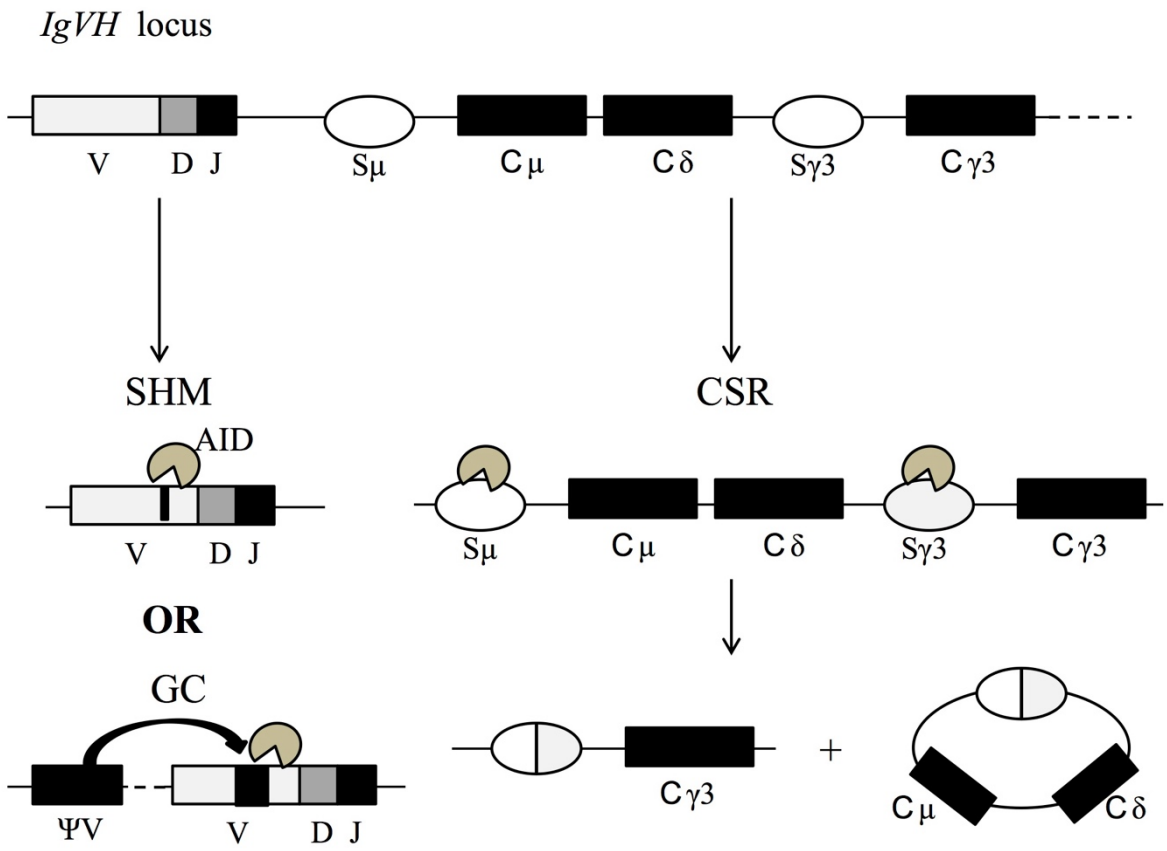


Figure 2. AID contributes to the diversification of the V region and the isotype switch of the C region.

AID is an essential factor for somatic hypermutation (SHM), gene conversion (GCV), and class switch recombination (CSR). The activity of AID induces SHM and GCV on the IgV region gene, and CSR on the S region.

Control Mechanisms of AID Expression and Activity

AID is 34 kDa protein that belongs to the APOBEC family, and acts specifically on the Ig locus and contributes to SHM, GCV, and CSR by changing dC into dU in DNA (6). AID is essential for the secondary diversification of the IgV region and the Ig constant region to improve antibody affinity and change effector functions, respectively. Deficiency in AID is a cause of hyper-IgM syndrome, which is one type of immunodeficiency diseases (7). On the other hand, AID triggers genetic damages when acting on genetic loci other than the immunoglobulin locus (8). Therefore, the activity of AID and subsequent gene mutating mechanisms are tightly regulated to specifically target the Ig locus.

The first mechanisms to regulate AID are based on the control of transcription and translation levels of AID through active transcriptional regulation (HoxC4, NFκB) (9) and post-transcriptional regulation (miR-155, miR-181) (10) (11).

The second mechanisms are through the subcellular localization of AID. AID is largely localized in the cytoplasm, and a part of AID shuttles between the nucleus and the cytoplasm (12). It has been shown that eEf1α keeps AID in the cytoplasm (13), and Hsp90 and Hsp40/DnaJa1 stabilize AID in the cytoplasm (14) (15). In addition, the C terminus of AID is involved in the Crm1-dependent nuclear export mechanism (16). It has been reported that cofactors regulate SHM and CSR activities by binding to AID. As a cofactor for SHM, mRNA nuclear export-associated factor GANP recruits AID to IgV (17). On the other hand, splicing-regulator PTBP2 and adaptor protein 14-3-3 recruit AID to S region. Transcriptional repressor KAP1-HP1 complex recruits AID to histone H3K9me3 on the S region gene (18) (19) (20), and PKA recruits AID to the S region gene by phosphorylation of AID at the Ser38 site (21). Additionally, splicing-related factor CTNNBL1, and transcription factor YY1 contribute accumulation of AID into the nucleus for CSR (22) (23) (Fig. 3).

The third mechanisms regulate the activity of AID. AID uses single-stranded DNAs

(ssDNA) as substrates that are generated in a transcription-dependent manner (24). It has been reported that in the IgV region gene, ssDNA patches are generated as substrates of AID by the transcription complex (25), whereas in the switch region, ssDNA stretches are formed as substrates of AID in long R-loop structures that depends on GC-rich sequences (26). It is also suggested that Spt5 that is associated with stalled RNA polymerase II provides target sites for AID (27). The Histone H3.3 promotes the formation of ssDNA by changing the genomic structure in the IgV region gene (28). In addition, the RNA exosome binds to AID and degrades nascent RNAs associated with the template DNA strand to expose both DNA strands as ssDNAs for AID (29) (Fig. 4). Also, the phosphorylation of AID contributes to association with cofactors and catalytic activation for SHM or CSR (30).

The fourth mechanisms regulate the protein degradation of AID (31). The REG-γ interacts with AID and accelerates AID degradation with proteasome (32).

As described above, it has been reported that AID is strictly regulated through a multi-layered process including the expression, localization, activity, and degradation of AID. However, there are many unclear points. For example, it is unknown what determines the transition of AID from the cytoplasm into the nucleus. However, most of AID co-factors that are reported to control SHM and CSR can target the entire genome, and thus it is unclear what determines the Ig specificity of AID. In addition, it remains uncovered what mechanism links between AID-co-factor-dependent regulatory mechanisms, which are involved in histone modification, transcription stall, and AID phosphorylation, etc. In order to elucidate mechanisms to control the activity of AID in SHM and CSR, it is necessary to clarify precise relationships between the functions of AID co-factors in AID-associating mechanisms.

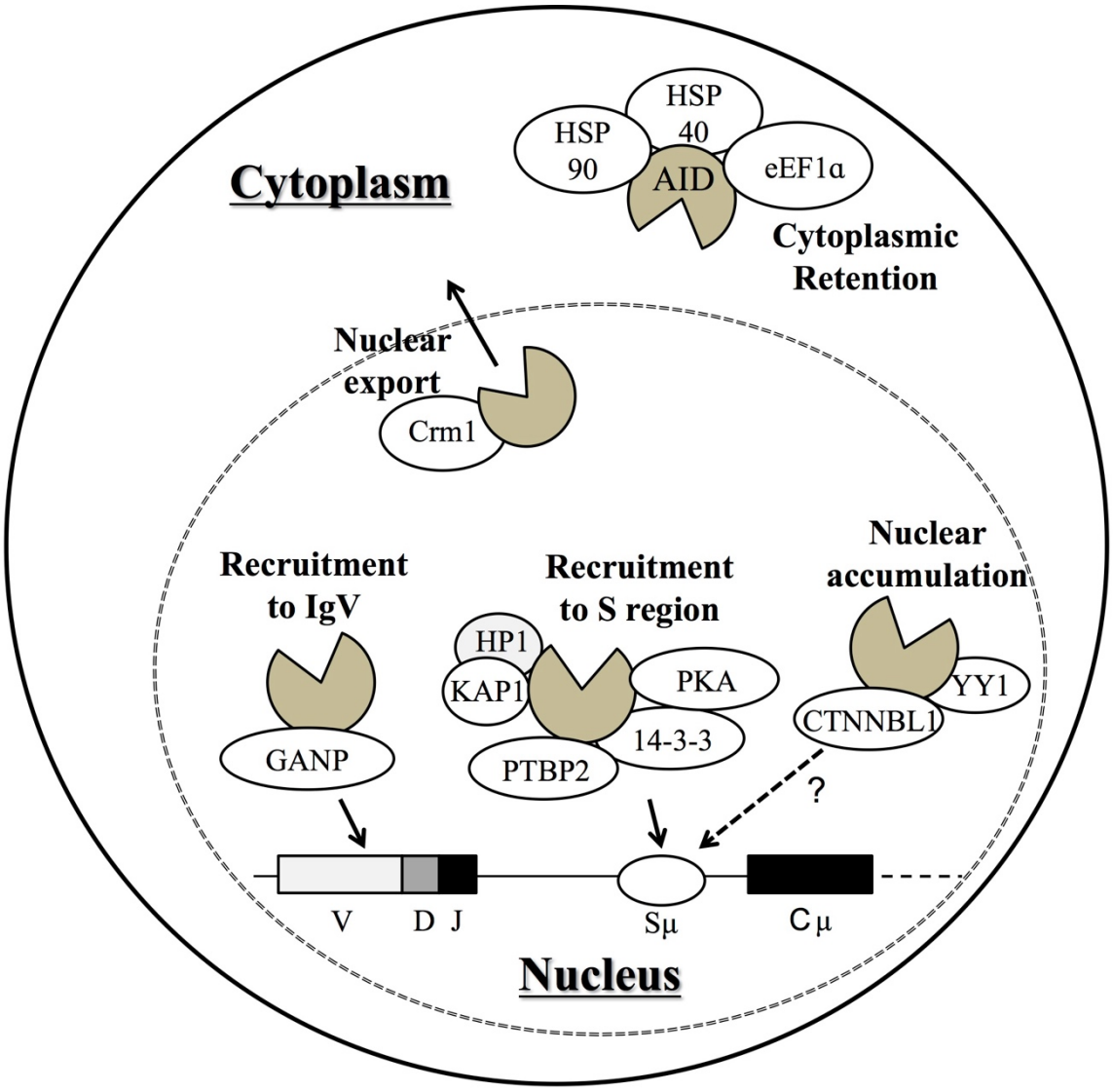


Figure 3. AID binging co-factor and its function.

It has been reported that AID interacts with co-factors for SHM, GCV, or CSR and is controlled in its localization. HSP90, HSP40, and eEf1α contribute to the stability and retention of AID in the cytoplasm. Crm1 exports AID outside the nucleus. PTBP2, 14-3-3, PKA, KAP1-HP1 complex, CTNNBL1, and YY1 regulate CSR. GANP regulates SHM.

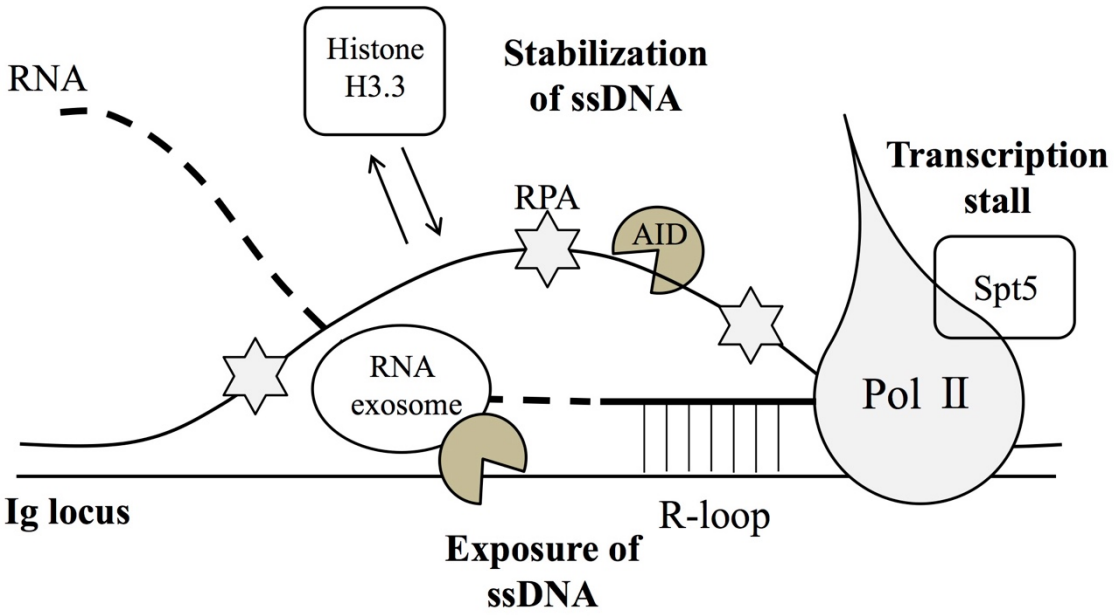


Figure 4. The formation of ssDNA on the Ig locus.

ssDNAs are substrates for AID. Spt5, RNA exosome, Histone H3.3, and RPA contribute to ssDNA formation.

Structure and Functions of SRSF1

The serine / arginine-rich protein (the SR protein) is an RNA binding protein, which has one or two RNA binding domains known as the RNA recognition motif (RRM) in the N-terminal region and one Arg-Ser rich domain (the RS domain) at the C-terminus. The SR proteins are important for constitutive and alternative splicing events of pre-mRNA, and in addition, are involved in a variety of cellular processes in RNA metabolism such as nuclear export and stabilization of mRNA, and maintenance of genome stability (33). In homo sapiens, 12 SR proteins have been reported, and SRSF1 is the prototypical SR protein (34). SRSF1 has two RRMs in the N-terminal region and one RS domain in the C-terminal region. It binds to exon splicing enhancers (ESS) of target pre-mRNAs on the RRMs, and promotes the accumulation of splicing-related proteins on the RS domain, thereby assisting the pre-mRNA splicing. SRSF1 deficiency causes R-loop formation throughout the genome and consequently induces DNA double stranded breaks (DSB) (35) (Fig. 5). In addition, the phosphorylation state of the RS domain regulates the shuttling of SRSF1 between the cytoplasm and the nucleus, and the localization into nuclear speckles (36).

The SRSF1-3 is an intron retention isoform, which is generated by intron inclusion: the intron between the exon3 and exon4 of the SRSF1 transcript is included to generate the SRSF1-3 mature mRNA and thereby encodes an alternate C-terminus instead of the RS domain (Fig. 5). Thus, SRSF1-3 might be a dominant negative form of SRSF1, because it has been reported to inhibit the splicing enhancing activity of SRSF1 *in vitro* (37).

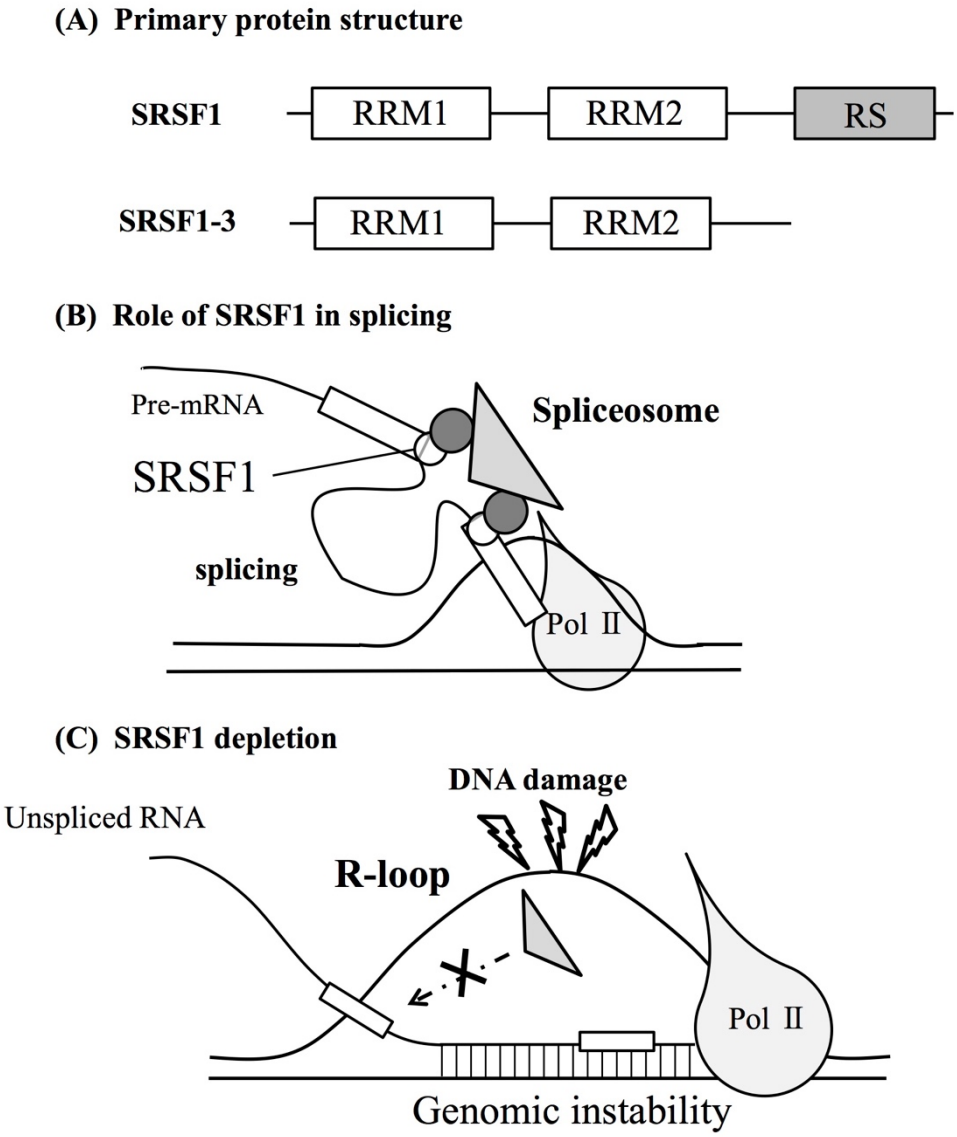


Figure 5. Structure and function of SRSF1.

SRSF1 (formerly ASF/SF2) has diverse roles in RNA metabolism, including splicing, nuclear export, stability control, regulation of translation, and maintenance of genomic stability. Depletion of SRSF1 causes genomic instability, resulting from the formation of transcription-coupled R-loops. (A) Primary structure of SRSF1 and SRSF1-3 protein. (B) Role of SRSF1 in RNA splicing. (C) SRSF1 depletion induces genomic instability through the formation of R-loops.

Aims of This Study

AID is an essential molecule for SHM, GCV and CSR, and the Ig locus specificity of AID might be controlled by known and unknown co-factors. However, detail mechanisms are unclear in many points.

SRSF1-3, a splicing isoform of SRSF1, was discovered as a novel co-factor that regulates SHM and GCV (38). Lack of SRSF1-3 expression causes deficiency in SHM and GCV, and re-introduction of SRSF1-3 into SRSF1-3-lacking cells restores SHM and GCV activities. However, the functional mechanisms of SRSF1-3 in the AID-dependent IgV diversification have not been clarified. In the present study, the author wishes to elucidate cellular and molecular mechanisms for SRSF1-3 in the IgV hypermutation. Detailed analyses of SRSF1-3 functions will also reveal functional interactions between AID and other co-factors, and will provide a foothold for understanding the control mechanisms of AID activity.

In the present study, the author revealed the following points.

SRSF1-3 contributes to the localization of AID into the nucleus.

It has been reported that the interaction of AID with transcription-coupled factors, such as splicing-related factors, CTNNBL1 (22) and PTBP2 (18), and RNA polymerase Ⅱ associated factors, Spt5 (27), RNA exosome subunits (29), and PAF complex (39), recruit AID to its target genes in the nucleus. SRSF1 is a nucleocytoplasmic shuttling protein, which contributes to a lot of RNA polymerase Ⅱ-coupled processes, and thus its splicing isoform, SRSF1-3, might alter SRSF1-assisting splicing and affect RNA polymerase Ⅱ-mediated transcriptional elongation on the IgV gene in the nucleus. Therefore, the author investigated whether SRSF1-3 recruits AID into the nucleus by interacting with AID.

Results of co-expression experiments of SRSF1-3 and AID in HEK293T cell line showed that AID localization was changed from the cytoplasm to the nucleus. SRSF1-3 was interacted and co-localized with AID. These results suggest that SRSF1-3 forms a protein complex with AID, and promotes the nuclear localization of AID for IgV hypermutation.

Overexpression of SRSF1-3 can enhance IgV hypermutation activity in hypermutating B cell lines, which constitutively express AID. However, when SRSF1-3 was overexpressed in a B cell line expressing a C-terminus-deleted AID mutant that lacks the nuclear export signal, the overexpression of SRSF1-3 did not affect IgV hypermutation, suggesting that functions of SRSF1-3, including the nuclear localization of AID, might be linked with the C terminus of AID.

References

(1) David G. Schatz, Patrick C. Swanson, V(D)J Recombination: Mechanisms of Initiation. Annu. Rev. Genet. 45, 167-202 (2011).

(2) Michael McHeyzer-Williams, Shinji Okitsu, Nathaniel Wang and Louise Mc Heyzer-Williams, Molecular programming of B cell memory. Nat. Rev. Immunol. 12, 24-34 (2012).

(3) Jayanta Chaudhuri, Ming Tian, Chan Khuong, Katrin Chua, Eric Pinaud and Frederick W. Alt, Transcription-targeted DNA deamination by the AID antibody diversification enzyme. Nature. 422, 726-730 (2003).

(4) Masamichi Muramatsu, Kazuo Kinoshita, Sidonia Fagarasan, Shuichi Yamada, Yoichi Shinkai, and Tasuku Honjo, Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell. 102, 553-563 (2000).

(5) Reuben S. Harris, Julian E. Sale, Svend K. Petersen-Mahrt, and Michael S. Neuberger, AID is Essential for Immunoglobulin V Gene Conversion in a Cultured B Cell Line. Current Biology. 12, 435-438 (2002)

(6) Silvestro G. Conticello, Cornelia J. F. Thomas, Svend K. Petersen-Mahrt, and Michael S. Neuberger, Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. Mol Biol Evol. 22, 367-377 (2005).

(7) Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R, Gennery A, Tezcan I, Ersoy F, Kayserili H, Ugazio AG, Brousse N, Muramatsu M, Notarangelo LD, Kinoshita K, Honjo T, Fischer A, Durandy A. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell. 102, 565-575 (2000)

(8) Adams JM, Gerondakis S, Webb E, Mitchell J, Bernard O, Cory S. Transcriptionally

active DNA region that rearranges frequently in murine lymphoid tumors. PNAS. 79, 6966-6970 (1982).

(9) Seok-Rae Park, Hong Zan, Zsuzsanna Pal, Jinsong Zhang, Ahmed Al-Qahtani, Egest J Pone, HoxC4 binds to the promoter of the cytidine deaminase AID gene to induce AID expression, classswitch DNA recombination and somatic hypermutation. Nat Immunol. 10, 540-550 (2009).

(10) de Yébenes VG, Belver L, Pisano DG, González S, Villasante A, Croce C, He L, Ramiro AR. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. JEM. 205, 2199-2206 (2008).

(11) Teng G, Hakimpour P, Landgraf P, Rice A, Tuschl T, Casellas R, Papavasiliou FN. MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. Immunity. 28, 621-629 (2008).

(12) Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M, Honjo T. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. PNAS. 101, 1975-1980 (2004)

(13) Häsler J, Rada C, Neuberger MS. Cytoplasmic activation-induced cytidine deaminase (AID) exists in stoichiometric complex with translation elongation factor 1α (eEF1A). PNAS. 108, 18366-18371 (2011).

(14) Orthwein A, Patenaude AM, Affar el B, Lamarre A, Young JC, Di Noia JM. Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp90. JEM. 207, 2751-2765 (2010).

(15) Orthwein A, Zahn A, Methot SP, Godin D, Conticello SG, Terada K, Di Noia JM. Optimal functional levels of activation-induced deaminase specifically require the Hsp40 DnaJa1. EMBO J. 31, 679-691 (2012).

(16) McBride KM, Barreto V, Ramiro AR, Stavropoulos P, Nussenzweig MC. Somatic

Hypermutation Is Limited by CRM1-dependent Nuclear Export of Activation-induced Deaminase. JEM. 199, 1235-1244 (2004).

(17) Maeda K, Singh SK, Eda K, Kitabatake M, Pham P, Goodman MF, Sakaguchi N. GANP-mediated recruitment of activation-induced cytidine deaminase to cell nuclei and to immunoglobulin variable region DNA. J Biol Chem. 285, 23945-23953 (2010).

(18) Nowak U, Matthews AJ, Zheng S, Chaudhuri J. The splicing regulator PTBP2 interacts with the cytidine deaminase AID and promotes binding of AID to switch-region DNA. Nat Immunol. 12, 160-166 (2010).

(19) Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, Thomas LM, Al-Qahtani A, White CA, Park SR, Steinacker P, Li Z, Yates J 3rd, Herron B, Otto M, Zan H, Fu H, Casali P. 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. Nat Struct Mol Biol. 17, 1124-1135 (2010).

(20) Jeevan-Raj BP, Robert I, Heyer V, Page A, Wang JH, Cammas F, Alt FW, Losson R, Reina-San-Martin B. Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. J Exp Med. 208, 1649-1660 (2011).

(21) Vuong BQ, Lee M, Kabir S, Irimia C, Macchiarulo S, McKnight GS, Chaudhuri J. Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. Nat Immunol. 10, 420-426 (2009).

(22) Conticello SG, Ganesh K, Xue K, Lu M, Rada C, Neuberger MS. Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNBL1. Mol Cell. 31, 474-484 (2008).

(23) Zaprazna K, Atchison ML. YY1 controls immunoglobulin class switch recombination and nuclear activation-induced deaminase levels. Mol Cell Biol. 32, 1542-1554 (2012).

(24) Bransteitter R, Pham P, Scharff MD, Goodman MF. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of

RNase. PNAS. 100, 4102-4107 (2003).

(25) Ronai D, Iglesias-Ussel MD, Fan M, Li Z, Martin A, Scharff MD. Detection of chromatin-associated single-stranded DNA in regions targeted for somatic hypermutation. JEM. 204, 181-190 (2007).

(26) Yu K, Chedin F, Hsieh CL, Wilson TE, Lieber MR. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. Nat Immunol. 4, 442-451 (2003).

(27) Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, Reina San-Martin B, Barreto V, Nieland TJ, Root DE, Casellas R, Nussenzweig MC. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. Cell. 143, 122-133 (2010).

(28) Romanello M, Schiavone D, Frey A, Sale JE. Histone H3.3 promotes IgV gene diversification by enhancing formation of AID-accessible single-stranded DNA. EMBO J. 35, 1452-1464 (2016).

(29) Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR, Januszyk K, Gregory RI, Deng H, Lima CD, Alt FW. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. Cell. 144, 353-363 (2011).

(30) Xu Z, Zan H, Pone EJ, Mai T, Casali P. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. Nat Rev Immunol. 12, 517-531 (2012).

(31) Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill JC, Reynaud CA. Proteasomal degradation restricts the nuclear lifespan of AID. JEM. 205, 1357-1368 (2008).

(32) Uchimura Y, Barton LF, Rada C, Neuberger MS. REG-γ associates with and modulates the abundance of nuclear activation-induced deaminase. JEM. 208, 2385-2391 (2011).

(33) Manley JL, Tacke R. SR proteins and splicing control. Genes Dev. 10, 1569-1579 (1996).

(34) Manley JL, Krainer AR. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev. 24, 1073-1074 (2010).

(35) Li X, Manley JL. Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. Cell. 122, 365-378 (2005).

(36) Cazalla D[1], Zhu J, Manche L, Huber E, Krainer AR, Cáceres JF. Nuclear export and retention signals in the RS domain of SR proteins. Mol Cell Biol. 22, 6871-6882 (2002).

(37) Zuo P, Manley JL. Functional domains of the human splicing factor ASF/SF2. EMBO J. 12, 4727-4737 (1993).

(38) Kanehiro Y, Todo K, Negishi M, Fukuoka J, Gan W, Hikasa T, Kaga Y, Takemoto M, Magari M, Li X, Manley JL, Ohmori H, Kanayama N. Activation-induced cytidine deaminase (AID)-dependent somatic hypermutation requires a splice isoform of the serine/arginine-rich (SR) protein SRSF1. PNAS. 109, 1216-1221 (2012).

(39) Willmann KL, Milosevic S, Pauklin S, Schmitz KM, Rangam G, Simon MT, Maslen S, Skehel M, Robert I, Heyer V, Schiavo E, Reina-San-Martin B, Petersen-Mahrt SK. A role for the RNA pol II-associated PAF complex in AID-induced immune diversification. JEM. 209, 2099-2111 (2012).

SRSF1-3 Contributes to Diversification of the Immunoglobulin Variable Region Gene by Promoting Accumulation of AID in the Nucleus

Abstract

Activation induced cytidine deaminase (AID) is essential for somatic diversification of the Ig variable region (*IgV*) in an *Ig* gene transcription-dependent manner. AID is excluded from the nucleus, where it normally functions, and consequently accumulated in the cytoplasm. However, the molecular mechanisms used for regulating AID localization remain to be elucidated. The SR-protein splicing factor SRSF1 is a transcription-coupled and nucleocytoplasmic shuttling protein, of which a splicing isoform, SRSF1-3, has been previously identified to play an essential role in *IgV* diversification in chicken DT40 cells. In this study, we examined whether SRSF1-3 contributes to *IgV* diversification by promoting nuclear localization of AID. AID expressed alone was localized predominantly in the cytoplasm. In contrast, co-expression of AID with SRSF1-3 leads to accumulation of AID along with SRSF1-3 in the nucleus and the formation of a protein complex that contained both AID and SRSF1-3, although SRSF1-3 was dispensable for nuclear import of AID. On the other hand, expression of either SRSF1-3, or a C-terminally-deleted AID mutant, which has been reported to lack cytoplasmic retention and nuclear export, increased *IgV* diversification in DT40 cells. However, it was found that overexpression of exogenous SRSF1-3 was not able to further enhance *IgV* diversification in DT40 cells expressing the truncated AID mutant. These results suggest that SRSF1-3 promotes nuclear localization of AID probably by forming a nuclear protein complex, which might stabilize nuclear AID and induce *IgV* diversification in an AID C-terminus-dependent manner.

Introduction

The immunoglobulin gene is diversified during B cell development by V(D)J recombination (1). When B cells encounter the specific antigen that binds to the B cell antigen receptor, B cells are activated and form germinal centers (2). Germinal center B cells express activation-induced cytidine deaminase (AID), which converts dC to dU on single-stranded (ss) DNAs to allow somatic hypermutation (SHM) on the Ig variable region genes (*IgV*) and class switch recombination (CSR) on Ig constant region genes (3,4). Since the off-target activity of AID can sometimes cause oncogenic mutations (5), the expression and activity of AID is precisely regulated through a multi-layered process involving regulation of AID expression by transcription factors (6) and miRNAs (7-9), protein phosphorylation (10-12), cytoplasmic retention (13,14), nuclear import (15) and export (16,17), protein degradation (18,19), targeting to the *Ig* locus (20-23), etc. It has been reported that AID is predominantly localized in the cytoplasm (16,17). Hsp90 and eEF1A play important roles in the cytoplasmic retention of AID (13,14). Crm1, which binds to the nuclear export signal at the AID C-terminus (16,17), promotes the nuclear export of AID to the cytoplasm. However, the molecular mechanisms that shuttle AID between the cytoplasm and the nucleus and that recruit AID to the *Ig* gene have not been fully elucidated.

The chicken B cell line DT40 constitutively expresses AID, and somatic hypermutation (SHM) and another *IgV*-diversifying process, gene conversion (GCV), both of which are initiated by a common intermediate, occur spontaneously in DT40 cells (24,25). We previously found an important role for a splice variant of the serine/arginine (SR)-rich protein splicing factor 1 (SRSF1), referred to here as SRSF1-3, in SHM and GCV (26). An engineered DT40 cell line, DT40-ASF, in which the endogenous SRSF1 gene is disrupted and the human SRSF1 cDNA is expressed (27), was unable to induce SHM and GCV, whereas introduction of SRSF1-3 into the DT40-ASF cells restored these *IgV* diversification events (26). The prototypical SR

protein, SRSF1 (formerly ASF/SF2 (28)), is a nucleocytoplasmic shuttling protein, which is not only essential for RNA splicing but also has a variety of roles in RNA metabolism, including nuclear export of mRNA, stability and quality control of mRNA, and translational regulation (29,30). SRSF1 is composed of two RRMs (RNA recognition motifs) in the N-terminal region and one serine-arginine-rich domain (RS domain) in the C-terminal region (29). The RRM domains are involved in specific binding to target pre-mRNAs and proteins (31). The RS domain is required for nucleocytoplasmic shuttling of SRSF1 (32), and acts a scaffold for assembly of splicing factors to stimulate splicing (33,34). The *SRSF1* genetic locus can generate several alternatively spliced transcripts (35,36), none of which have not been reported to have a cellular function except for SRSF1-3 (26).

SRSF1-3 is one alternatively spliced product of SRSF1, which shares two RRMs with SRSF1 but contains an alternate C-terminus that is generated by intron inclusion (Fig. 1A). The alternate C-terminus is not conserved between species, lacks the RS domain, and does not include any other characteristic protein motifs that indicate its function. Thus, the molecular role of SRSF1-3 in *IgV* diversification remains unknown, although SRSF1-3 has been reported to inhibit the splicing-enhancing function of SRSF1 *in vitro* (37). Recently, transcription-coupled factors, such as CTNNBL1 (38), Spt5 (20), PTBP2 (22), RNA exosome subunits (39), and the PAF complex (40) have been shown to associate with AID, thereby forming a protein complex that induces SHM and CSR. It has also been reported that AID co-localizes with splicing factors in subnuclear domains where SRSF1 is predominantly enriched (41). The increase in the number of reports connecting AID and SRSF1 prompted us to explore whether SRSF1-3, an isoform of SRSF1, contributes to the localization of AID in the nucleus. In this study, we investigated a role for SRSF1-3 in the subcellular localization of AID using fluorescently tagged fusion proteins. The results suggest that SRSF1-3 promotes accumulation of AID in the nucleus during *IgV* diversification.

Material and Methods

Cell lines

Wild-type DT40 cells and 293T cells were obtained from the RIKEN Cell Bank. DT40-ASF cells (27), DT40-SW cells (45), and DT40-SWΔC cells (44) were described previously. DT40 cells were cultured in RPMI-1640 medium (Nissui Pharmaceutical, Tokyo, Japan) supplemented with 10% fetal bovine serum (Nichirei Biosciences, Tokyo, Japan), 1% chicken serum (Nippon Bio-Test Laboratories, Tokyo, Japan), 50 μM 2-mercaptoethanol, 2 mM glutamine, 1 mM pyruvic acid, 100 μg/mL penicillin G, and 50 μg/mL streptomycin at 40°C in 5% CO_2. 293T cells were cultured in Dulbecco's modified Eagle's medium (Life Technologies, Gaithersburg, MD, USA) supplemented with 10% fetal bovine serum and 50 μM 2-mercaptoethanol at 37°C in 5% CO_2

Construction of expression vectors and transfection

Expression vectors encoding mCherry-SRSF1-3 and AID-GFP fusion proteins were constructed as follows. The chicken SRSF1-3 expression vector, in which a FLAG-tagged chicken *SRSF1-3* cDNA fragment was inserted into the pCI-bsr vector (46), has been described previously (26). FLAG-tagged human *SRSF1-3* cDNA fragment was amplified by PCR from a human Ramos B cell line cDNA library, which was prepared from total RNA of Ramos cells by reverse transcription, using a sense primer, 5'-GGAATTCGCCACCATGGACTACAAGGACGACGACGACAAGGCCATGTCGGGAGGTGGTGTGATTCG-3', and an antisense primer, 5'-CCTTGAAATTCCACTGTTAAGACC-3', and cloned into the pCI-bsr vector as described previously (26). A DNA fragment encoding mCherry was excised from the pmCherry-C1 vector (Clontech Laboratories, Mountain View, CA, USA) after the *Age* I site of the vector was converted to an *Xho* I site with an oligonucleotide linker (5'-CCGGCTCGAGA-3'), and inserted at the *Xho* I sites of the FLAG-

tagged chicken and human SRSF1-3 expression vectors, in frame, just upstream of the FLAG-tagged SRSF1-3s to yield plasmid expression vectors encoding mCherry-tagged chicken and human SRSF1-3s, respectively. A Myc-tagged chicken AID (Myc-AID) DNA fragment was generated from a chicken *AID* cDNA clone (45) by PCR using a sense primer, 5'-GGAATTCGCCACCATGGAGCAGAAACTCATCTCTGAAGAGGATCTGGGCATGGA CAGCCT-3', and an antisense primer, 5'-CTCCTTTCTTGGCTGGGTGAGAGGTCCATA-3'. For constructing AID-GFP, another antisense primer, 5'-CATCTGAGAGAGAATTCAGCTGACATGGAC-3', was used for PCR. The DNA fragment for construction of AID-GFP was fused in frame with a DNA fragment encoding EGFP by inserting between the *Eco*R I and *Kpn* I sites in pEGFP-N1 vector (Clontech Laboratories). The DNA fragment encoding Myc-AID and the fused DNA fragment encoding AID-GFP were inserted between the *Eco*R I and *Not* I sites in the pCI-bsr vector to yield plasmid expression vectors encoding Myc-AID and AID-GFP, respectively. PCR was carried out using KOD FX neo DNA polymerase (Toyobo, Osaka, Japan), and the nucleotide sequences of all PCR-cloned fragments were confirmed by sequencing analysis.

For microscopic analyses, 293T cells were seeded onto glass cover slips that had been coated with 0.1% gelatin, and for biochemical analyses 293T cells were on culture dishes. Cells were cultured to 70-80% confluency and then were transiently transfected with expression vectors using the X-treme GENE HP DNA transfection reagent (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's instructions. In some experiments, the pCI-bsr vector was used as an empty vector for mock transfection. Cells were analyzed by confocal microscopy and biochemical analyses 36 h after transfection. DT40 cells were transfected with expression vectors by electroporation using the Neon Transfection System (Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA). After selection of drug-resistant clones in a medium containing 20 μg/mL Blasticidin S (Kaken Pharmaceutical, Tokyo, Japan), expression of

mCherry-cSRSF1-3 and AID-GFP were evaluated by flow cytometry.

Microscopic analysis

Prior to microscopic analysis for AID-GFP and mCherry-SRSF1-3, transfected 293T cells on cover slips were fixed with 4% paraformaldehyde for 15 min, permeabilized in 0.1% Triton X-100/PBS, and mounted on slide glasses with ProLong® Gold Antifade Reagent with DAPI (Invitrogen). DT40 cells were similarly fixed with 4% paraformaldehyde, immersed in 0.1% Triton X-100/PBS, and mounted on glass slides. In some experiments, DT40 cells were treated with 10 μM MG132 (Sigma-Aldrich, St. Louis, MO, USA), and/or 10ng/mL leptomycin B (Cayman Chemical, Ann Arbor, MI, USA) for either 6 h or 2 h, respectively, or for indicated time periods, before analysis.

For immunofluorescent staining, DT40 cells were fixed in 1% paraformaldehyde for 30 min, and permeabilized with 0.5% nonidet P-40. After blocking in 2% goat serum, cells were incubated with mouse anti-FLAG mAb (clone M2; Sigma-Aldrich) followed by Alexa Fluor 594-conjugated rabbit anti-mouse IgG (Molecular Probes, Eugene, OR, USA).

Cell images were acquired using an FV1000 confocal laser-scanning microscope and FluoView software (Olympus, Tokyo, Japan) at room temperature. Image capturing was carried out using objective lenses, UPlanSApo 60× or UPlanFLN 40× (Olympus) with immersion oil. Cells that showed bright GFP and/or mCherry fluorescence were selected for analysis in each microscopic view under conditions where background fluorescence was undetectable in each detection channel and where the gain of each detector and the power of each laser were adjusted to minimize crosstalk detection between channels. The mean fluorescence intensities (MFIs) of AID-GFP in the nuclear and cytoplasmic regions of interest in chosen cells were quantified using ImageJ software (62). Regions showing green or blue fluorescence, which corresponded to AID-GFP or DNA, were considered as representing the whole cell or the nucleus of each

cell, respectively. The MFI of AID-GFP in the cytoplasmic region of each cell was calculated by subtracting that in the nuclear region from that in the whole cell region. The subcellular localization of AID-GFP was evaluated by calculating the ratio of the MFI of AID-GFP in the nuclear region over that in the cytoplasmic region. Cells having more nuclear AID-GFP showed higher ratio values. When the subcellular localization of mCherry-cSRSF1-3 was evaluated, images by differential interference contrast microscopy were used to determine the whole cell region of each cell.

Quantitative RT-PCR

The transcription level of the *SRSF1-3* gene was determined by quantitative RT-PCR, using Thunderbird SYBR qPCR Mix (Toyobo) and the iQ5 real-time PCR detection system (Bio-Rad Laboratories, Hercules, CA), as described previously (26). For each sample, the cycle threshold (Ct) values for the *SRSF1-3* transcript were determined, and data were normalized to the Ct value for the *GAPDH* transcript.

Subcellular fractionation and western blotting

The nuclear and cytoplasmic fractions were separated with the NE-PER Nuclear and Cytoplasmic Extraction Kit (Thermo Fisher Scientific Pierce Biotechnology, Rockford, IL, USA). The nuclear and cytoplasmic fractions derived from cells cultured on a 5-cm dish for each sample were finally adjusted in volume up to 200 μL with MilliQ water (Millipore, Billerica, MA, USA), and mixed with the equal volume of 2 × Laemmli sample buffer. For western blotting, whole cell lysates were prepared from 293T cells cultured on a 10-cm dish for each sample by suspending in 300 μL of Laemmli sample buffer, and by sonicating cells with a Bioruptor UCD-250 (Cosmo Bio, Tokyo, Japan). Following these, aliquots of cell fractions (15 μL each) or whole cell lysates (10 μL each) were subjected to SDS-PAGE (12.5%

acrylamide) and transfer onto a polyvinylidene difluoride membrane (Immobilon P; Millipore), and probed with one of the following antibodies; biotinylated goat anti-GFP antibody (ab6658; Cambridge, MA, USA), rabbit anti-mCherry antibody (GTX128508; Genetex, Irvine, CA, USA), mouse anti-ß-actin mAb (sc-47778; Santa Cruz Biotechnology, Santa Cruz, CA, USA), mouse anti-AID mAb (L7E7; Cell Signaling Technology, Beverly, MA, USA), mouse anti-FLAG mAb, goat anti-Hsp90 antibody (AF3775; R&D Systems, Minneapolis, MN, USA), rabbit anti-acetylhistone H4 antibody (06-866; Millipore), or mouse anti-c-Myc mAb (clone 4A6; Millipore), followed by detection of immunoreactive proteins using one of the following secondary reagents as appropriate; HRP-conjugated-streptavidin (LRPN1231V; GE Healthcare, London, UK), anti-rabbit IgG-HRP (NA934V; GE Healthcare), anti-mouse IgG-HRP (NA931V; GE Healthcare), or donkey anti-goat IgG-HRP (sc-2020; Santa Cruz Biotechnology). The mAb for mouse and human AIDs was able to detect the chicken ortholog as previously described (44). Blots were developed using Western Lightning Plus-ECL (PerkinElmer, Waltham, MA, USA), and chemiluminescence was detected using the ChemiDoc XRS system (Bio-Rad Laboratories) according to the manufactures' instructions. The intensity of each signal was quantified with Quantity One software (Bio-Rad Laboratories). The nuclear localization of AID was evaluated by calculating the ratio of the nuclear AID signal over the cytoplasmic AID signal.

Immunoprecipitation

293T cells were collected from a 10-cm dish for each sample 36 h after transfection with expression vectors encoding Myc-AID and FLAG-cSRSF1-3, and whole cell lysates were prepared with a Bioruptor sonicator in 450 μL of a lysis buffer [5 mM Tris-HCl pH 7.5, 10% (vol/vol) glycerol, 0.1% (vol/vol) Triton X-100, 150 mM NaCl] with a protease inhibitor mixture (Nakalai Tesque, Kyoto, Japan). Cell free extracts were prepared by centrifugation of

whole cell lysates at 1,500g for 5 min. An aliquot of cell free extracts (50 μL) was mixed with the equal volume of 2 × Laemmli sample buffer and used as an input control. Anti-FLAG mAb was adsorbed on Dynabeads™ Protein G magnetic beads (Invitrogen) using a 30-min incubation at room temperature, and the mAb-immobilized magnetic beads were then incubated with the rest of cell free extracts (approximately 400 μL) at 4°C for 1 h. After washing the beads, proteins bound to the beads were suspended in 50 μL of Laemmli sample buffer and used as immunoprecipitated samples. Aliquots of inputs and immunoprecipitates (10 μL each) were analyzed by SDS-PAGE and western blotting. For western blotting after immunoprecipitation, Mouse Trueblot Ultra anti-mouse Ig HRP (18-8817-31; Rockland, Gilbertsville, PA, USA) was used to avoid detection of the anti-FLAG mAb that was used for immunoprecipitation, and Western Lightning Pro ECL (PerkinElmer) was used to detect signals at a higher sensitivity.

Analysis of IgV diversification

The SRSF1-3 expression vector was introduced into DT40-SW and DT40-SWΔC cells, which express the wild-type AID and a C-terminally-deleted AID, respectively (44,45), as described previously (26). AID expression in both of these cell lines can be reversibly switched on or off by treating with 4-hydroxytamoxyfen (4-OHT) (45). GCV frequency in these cell lines was evaluated using a G/B construct as an artificial GCV substrate, as described previously (26,46). The G/B construct is composed of a CMV promoter-driven enhanced blue fluorescent protein (*EBFP*) gene and an un-transcribed enhanced green fluorescent protein (*EGFP*) gene (Fig. 5B). When GCV occurs in the *EBFP* gene using the *EGFP* gene as a sequence donor, the original blue fluorescence of the G/B construct-bearing cells is converted to a strong green fluorescence. The G/B construct was integrated into the un-rearranged *IgVL* allele in DT40 cells overexpressing SRSF1-3 as previously described (46). After treatment with 4-OHT, single

AID-expressing cells, which can be detected based on GFP expression linked with AID through an IRES were sorted as strong blue (GCV-un-induced) and weak green (AID-expressing) fluorescent cells into 96-well plates using a FACS Aria equipped with Auto Cell Deposit Unit (BD Biosciences, San Diego, CA, USA). Colonies were transferred to 24-well plates and maintained for weeks at cell densities between approximately 1×10^5 and 2×10^6 cells. The frequency of cells developing a strong green fluorescence signal was analyzed using FACS Array (BD Biosciences).

The nucleotide sequence analysis of the *IgVH* gene was carried out as described previously (26). DT40-SWΔC cells with or without overexpression of SRSF1-3 were cultured with AID expression being on for 3 weeks. The *IgVH* gene was amplified from the genomic DNA of the cultured cells by PCR with a primer pair specific to the chicken *IgVH* gene (26), then cloned and sequenced. Mutation events including GCV and SHM were identified by comparing mutated sequences with the original un-mutated sequence as described previously (44). The frequency of mutations was calculated by dividing the number of all mutations by then number of total analyzed nucleotides.

Statistical analyses

Statistical analyses were carried out using GraphPad Prism software (GraphPad Software, La Jolla, CA, USA). The unpaired two-tailed Student's t test was performed to evaluate statistically significant differences between two groups. The Pearson correlation analysis was performed to calculate correlation coefficients. The chi-square test was performed to evaluate difference in the mutation frequency of the *IgV* gene between in the presence and absence of overexpressed SRSF1-3. Probability (p) values are indicated in the Figure legends.

Results

Co-expression of AID with SRSF1-3 Promotes Accumulation of AID in the Nucleus

To visualize the subcellular localization of AID and SRSF1-3, chicken AID (tagged with GFP at the C-terminus; AID-GFP) and chicken and human SRSF1-3s (tagged with mCherry at the N-terminus; mCherry-cSRSF1-3 and mCherry-hSRSF1-3) were transiently expressed in human 293T cells. DT40 cells stably expressing AID-GFP, mCherry-cSRSF1-3, or FLAG-tagged chicken SRSF1-3 (FLAG-cSRSF1-3) were established by transfection and subsequent drug-selection. The subcellular localization of these fluorescent fusion proteins in these different cells was then observed by confocal microscopy. As described previously, when AID-GFP was expressed alone, i.e., without exogenous expression of SRSF1-3, AID-GFP was found predominantly in the cytoplasm in both 293T cells and DT40 cells (Fig. 1B and 1C, upper panels, respectively) (16,17,42), indicating that most of the AID protein is excluded from the nucleus in a Crm1-dependent manner (17). In 293T cells, although detectable levels of mCherry-cSRSF1-3 were found in the nucleus, mCherry-cSRSF1-3 was found to be mainly localized in the cytoplasm (Fig. 1B, middle panels) whereas the human ortholog (mCherry-hSRSF1-3) was localized predominantly in the nucleus (Fig. 1B, lower panels). In addition, when mCherry-cSRSF1-3 or FLAG-cSRSF1-3 were stably expressed in DT40 cells, the exogenously expressed SRSF1-3 was detected in the nucleus by mCherry fluorescence or by immunofluorescent staining, respectively (Fig. 1C). It should be noted that the levels of mCherry-cSRSF1-3 and FLAG-cSRSF1-3 as well as the frequency of cells expressing a detectable level of these tagged cSRSF1-3s were very low in DT40 cells (Fig.1C, and 1D, upper panel), and these detected mCherry-cSRSF1-3 and FLAG-cSRSF1-3 were localized predominantly in the nucleus in DT40 cells (Fig. 1D, lower panel). This low frequency of detectable tagged cSRSF1-3-expressing cells was confirmed by flow cytometry (Supplemental Fig. S1). These results suggest that SRSF1-3 might be expressed at a low level and is localized

in the nucleus under physiological conditions.

The cytoplasmic accumulation of mCherry-cSRSF1-3 in 293T cells might be due to an unstable conformation of the mCherry and chicken SRSF1-3 fusion protein and/or an unphysiologically overexpressing condition, but we considered that the nuclear mCherry-cSRSF1-3 in 293T cells was functional, according to the following observations: the nuclear localization of mCherry-cSRSF1-3 was localized in the nucleus in the physiological host DT40 cells (Fig. 1C), the human orthlog was also localized in the nucleus in 293T cells (Fig. 1B), and importantly, mCherry-cSRSF1-3 was able to compensate the *IgV* hypermutation-inducing activity of SRSF1-3 in SRSF1-3-deficient DT40-ASF cells (Supplemental Fig. S2). Thus, we next applied the experimental system using 293T cells and fusion proteins for co-expression experiments.

Interestingly, when AID-GFP was co-expressed with mCherry-cSRSF1-3 in 293T cells, AID-GFP levels increased in the nuclear compartment (Fig. 2A). The proportion of AID-GFP localized in the nucleus was significantly higher in cells co-expressing AID-GFP and mCherry-cSRSF1-3 than in cells expressing AID-GFP alone (Fig. 2B).

In addition, AID-GFP nuclear localization was positively correlated with the MFI level of co-expressed mCherry-cSRSF1-3 (Fig. 2C), whereas it was not correlated with the MFI level of AID-GFP when AID-GFP was expressed alone (Fig. 2D). Human SRSF1-3 had the same effect on AID localization as the chicken counterpart (Fig. 2A and 2B), suggesting that avian and mammalian SRSF1-3s play a similar role in the subcellular localization of AID. It should be noted that AID tended to be co-localized with SRSF1-3 in both the nucleus and the cytoplasm (Fig. 2A). Similarly, the subcellular localization of AID in the nucleus was also observed when AID-GFP and mCherry-SRSF1-3 were co-expressed in another adherent cell line, mouse NIH3T3 cells (data not shown). To ensure that the accumulation of green fluorescence in the nucleus resulted from altered localization of the AID-GFP fusion protein and not from diffusion

of GFP fragments derived from degradation of the AID-GFP fusion protein, the fusion proteins expressed in 293T cells were analyzed by western blotting. The results showed that co-expression with mCherry-cSRSF1-3 did not result in degradation of the AID-GFP fusion protein (Fig. 3A). Furthermore, subcellular fractionation analysis of 293T cells transiently expressing Myc-tagged AID (Myc-AID) and FLAG-tagged chicken SRSF1-3 (FLAG-cSRSF1-3) without fluorescent protein tags also confirmed that co-expression of AID with SRSF1-3 significantly enhanced accumulation of AID in the nuclear fraction in a fluorescent tag-independent manner (Fig. 3B). Taken together, the results from these cell biological and biochemical analyses reveal that SRSF1-3 promotes accumulation of AID in the nucleus.

We next wished to explore whether SRSF1-3 contributes to the accumulation of AID in the nucleus through interaction with AID as is seen for other transcription-coupled factors, or through an indirect effect. The interaction between AID and SRSF1-3 was examined by co-immunoprecipitation experiments. When Myc-AID was co-expressed with FLAG-cSRSF1-3 in 293T cells, Myc-AID was co-immunoprecipitated with FLAG-cSRSF1-3 using an anti-FLAG mAb from cell lysates (Fig. 3C), suggesting that AID and SRSF1-3 are found in the same protein complex.

SRSF1-3 is Dispensable for Import of AID into the Nucleus

It has been reported that AID proteins imported from the cytoplasm into the nucleus are immediately re-exported or degraded (15,18). Thus, at steady state, AID proteins appear to be mostly in the cytoplasm. We next examined whether SRSF1-3 could promote the import of AID into the nucleus. We generated cell lines stably expressing the AID-GFP fusion protein in SRSF1-3-sufficient wild-type DT40 cells and SRSF1-3-deficient DT40-ASF cells, which are competent and incompetent in *IgV* diversification, respectively (26). In contrast to the results of the transient co-expression experiments in 293T cells (Fig. 2), AID was observed

predominantly in the cytoplasm either in the presence or absence of SRSF1-3 expression in these DT40 lines (Fig. 4A, upper left panels). This is probably due to the immediate degradation and/or nuclear re-export of the imported AID. When proteasomal protein degradation or protein export from the nucleus were inhibited by treating these cells with specific inhibitors, MG132 or leptomycin B, respectively, the treatment with either of MG132 or leptomycin B resulted in the increased accumulation of AID in the nucleus, both in the presence and absence of SRSF1-3 (Fig. 4A, lower panels, and 4B). The inhibition of both proteasomal protein degradation and protein nuclear export further increased the nuclear accumulation of AID, indicating that proteasomal protein degradation and protein nuclear export independently exclude AID from the nucleus (Fig. 4A, upper right panels, and 4B). Interestingly, when only proteasomal protein degradation was inhibited by MG132 treatment, the nuclear accumulation of AID was lower in SRSF1-3-sufficient wild-type DT40 cells than SRSF1-3-deficient DT40-ASF cells (Fig. 4B). This difference was confirmed by a time course experiment of MG132 treatment (Fig. 4C). Taken together, these results suggest that SRSF1-3 is dispensable for the nuclear import of AID but may be involved in the nuclear accumulation of AID through other mechanisms.

A role for SRSF1-3 in IgV Diversification Might Depend on the C-terminus of AID

The C-terminus of AID is involved in several molecular processes during AID-dependent *IgV*diversification. For example, approximately 10 amino acids of the C-terminus encode a nuclear export signal (NES), so that deletion of the C-terminus leads to not only accumulation of AID proteins in the nucleus (16,17), but also enhanced SHM and GCV on the *IgV* gene (43,44). Similarly, overexpression of SRSF1-3 can promote nuclear accumulation of AID (Fig. 2) and in addition enhance SHM and GCV frequency in DT40 cells (26). Thus, we examined whether SRSF1-3 could contribute to SHM and GCV cooperatively or independently through functions associated with the AID C-terminus.

DT40 cell lines that express the wild-type chicken AID and a C-terminally-deleted (10 amino acids) mutant (AIDΔC) under the control of an exogenous CAG promoter, respectively, were previously established (44,45). An SRSF1-3 overexpression vector was introduced into these cell lines, which expresses sufficient endogenous SRSF1-3 for SHM and GCV to occur spontaneously, and stable SRSF1-3-overexpressing lines were established (Fig. 5A). An artificial DNA substrate encoding fluorescent proteins was introduced into the *IgVL* locus of DT40 cells. This substrate contained the gene encoding BFP (blue fluorescent protein) under the control of the CMV promoter. Upstream of this was a promoter-less DNA sequence encoding GFP. After GCV occurs, using GFP as a template, it results in mutation of the BFP gene at a single base (199T to 199C) to change it from encoding BFP to encoding GFP. This method can therefore efficiently detect and estimate AID-dependent GCV events through the generation of a cell population that has a high level of GFP expression (46) (Fig. 5B). Cells in which GCV occurred were observed during flow cytometry as GFP^{hi} cells. The results showed that in the wild-type AID-expressing control cells GFP^{hi} cells were observed (Fig. 5C), but were not observed in cells without AID expression (Fig. 5D) (46). When wild-type AID was expressed in DT40 cells, overexpression of SRSF1-3 resulted in a significantly higher GCV frequency compared with the control cells without SRSF1-3 overexpression (Fig. 5C). DT40 cells expressing AIDΔC similarly showed a higher GCV frequency than the wild-type control as described previously (44). In contrast, overexpression of SRSF1-3 in AIDΔC –expressing cells was not able to induce a further increase in GCV frequency. This analysis using the artificial substrate was considered to be quantitative within the range of GCV frequency measured but it did not reach a plateau, because when AIDΔC expressing cells, which already showed a high GCV frequency, were cultured for extended periods, more GFP positive cells were able to be detected over time (Fig. 5D).

To confirm the results using the artificial GCV substrate, nucleotide sequence analysis of the

Ig heavy chain variable region (*IgVH*) gene was carried out (Fig. 5E). The *IgVH* genes were cloned from the wild-type AID or AIDΔC-expressing DT40 cells, with or without overexpression of SRSF1-3, after 3 weeks of culture and sequenced. A significant number of mutations accumulated in the *IgVH* gene in AIDΔC-expressing DT40 cells that did not have overexpression of SRSF1-3 as reported previously (44) (Fig. 5D, left panel). The mutation frequency in AIDΔC-expressing DT40 cells (6.9×10^{-3} mutations/bp) was higher than the value the wild-type AID-expressing DT40 cells (0.98×10^{-3} mutations/bp). However, overexpression of SRSF1-3 did not affect the mutation frequency of the *IgVH* gene in AIDΔC-expressing DT40 cells, whereas that resulted in an obvious increase of the mutation frequency in the wild-type AID-expressing DT40 cells as reported previously (26) although the difference in the mutation frequency between cells with and without overexpression of SRSF1-3 in this experiment was not statistically significant by the chi-square test (Fig. 5E). These data reveal that a role for SRSF1-3 in *IgV* diversification may be genetically upstream of that of the AID C-terminus, or that SRSF1-3 may function cooperatively through an AID C-terminus-dependent mechanism for *IgV* diversification.

Discussion

To guarantee that SHM occurs specifically on the *Ig* locus in activated B cells, AID expression and activity are tightly and spatiotemporally regulated at multiple levels, including nuclear import, accumulation, export, and degradation. Although the nucleus is the site where AID exerts its cellular functions, AID is predominantly retained in the cytoplasm and excluded from the nucleus by protein export and degradation. In the present study, overexpression of SRSF1-3 revealed that SRSF1-3 may induce nuclear accumulation of AID and promote *IgV* diversification in an AID C-terminus-dependent manner. SRSF1-3 formed a protein complex with AID as has been reported for other transcription-coupled factors that are responsible for SHM and CSR (20,22,38-40), and furthermore SRSF1-3 and AID have been previously shown to be recruited to the *IgV* gene (26). Thus, we propose that SRSF1-3 may be involved in the formation and maintenance of a stable transcription-coupled protein complex responsible for inducing *IgV* gene diversification in the nucleus.

The present experiments using fluorescently tagged proteins indicated that the AID-GFP fusion protein appeared to be excluded from the nucleus in 293T cells and DT40 cells that might express a low level of endogenous SRSF1-3 (Fig. 1B and 1C) (26). The nuclear exclusion of AID is due to nuclear export and protein degradation, since this localization was sensitive to leptomycin B and MG132 treatment. Although endogenous SRSF1-3 may contribute to the accumulation of AID, the level of the AID protein is kept very low in the nucleus at steady state, even in the presence of physiological levels of SRSF1-3, partially because the level of endogenous SRSF1-3 is very low in DT40 cells (26). In addition, particularly in hyper-mutating cells such as DT40 cells and human Ramos cells, AID is excluded from the nucleus even in the absence of nuclear export of AID because of accelerated proteasomal degradation (18,44). This proteasomal degradation of AID in these hypermutating cells may be induced in an AID deaminase activity-dependent manner (19). Overexpressed SRSF1-3 may manifest a role for

endogenous SRSF1-3 in nuclear accumulation of AID by competing with intra-nuclear processes, including nuclear export and proteasomal degradation of AID.

SRSF1-3 tagged with mCherry or FLAG was observed in the nucleus in either 293T cells or DT40 cells (Figs. 1-3). SRSF1 is a nucleocytoplasmic shuttling protein, which moves across the nuclear membrane pore, along with a number of other proteins and mRNAs, as large ribonucleoprotein complexes (29,30). Because the interaction between RRM2 and RS domains regulates the phosphorylation-dependent nuclear import of SRSF1 (47), SRSF1-3, which lacks the RS domain, can be imported into the nucleus in the absence of phosphorylation-dependent regulation. Indeed it has been previously shown that the RS domain is not required for nuclear import of SRSF1 (48). Thus, it could be presumed that SRSF1-3 might contribute to the nuclear import of AID as a transportation carrier through either a direct or indirect interaction with AID (Fig. 3C). However, this is unlikely because AID was able to localize to the nucleus even in the absence of SRSF1-3 (Fig. 4B), although a supportive role for SRSF1-3 in nuclear import cannot be excluded. In addition, the AID molecule has its own nuclear import signal, which enables AID to move into the nucleus through an importin-mediated mechanism (16,49).

Whereas expression of either SRSF1-3 or AIDΔC was able to increase the nuclear localization of AID and the frequency of SHM and GCV, co-expression of both molecules did not show any additive effects on SHM and GCV (Fig. 5C and 5D). Thus, SRSF1-3 might function cooperatively with AID at the C-terminus of AID. It has been reported that deletion of the AID C-terminus results in decreased CSR but intact SHM, implying that the C-terminus of AID might have a role in determining which mechanism, either CSR or SHM, AID is engaged in (16,17,43,44,50). Crm1-dependent nuclear export has been well characterized as a function of the AID C-terminus (17). However, some AID C-terminal mutants, which are deficient in CSR but probably sufficient for SHM, showed intact Crm1-mediated nuclear export (15,51,52), and conversely inhibition of Crm1 by leptomycin B did not affect CSR (53). These recent

reports suggest that Crm1-dependent nuclear export might not contribute to a mechanism that regulates the choice between CSR and SHM, and therefore this is unlikely to be the function of the AID C-terminus that SRSF1-3 cooperates with to induce *IgV* diversification in parallel with the nuclear accumulation of AID. The AID C-terminus has been also reported to be involved in several other molecular functions of AID, such as binding to mRNA (54), dimerization (55), cytoplasmic retention (14,15), recruitment of the non-homologous end joining DNA repairing mechanisms on the *Ig* gene (56-61), etc., and most of these are required for CSR. Further investigations into the roles of SRSF1-3 will reveal whether SRSF1-3 contributes to the regulation of these functions listed above or to the regulation of other intranuclear processes to induce AID-dependent *IgV* diversification.

References

(1) Bassing, C. H., Swat, W., and Alt, F. W. (2002) The mechanism and regulation of chromosomal V(D)J recombination. *Cell* **109**, S45-55

(2) Victora, G. D., and Nussenzweig, M. C. (2012) Germinal centers. *Annu. Rev. Immunol.* **30**, 429-457

(3) Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553-563

(4) Chaudhuri, J., and Alt, F. W. (2004) Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat. Rev. Immunol.* **4**, 541-552

(5) Ramiro, A. R., Jankovic, M., Eisenreich, T., Difilippantonio, S., Chen-Kiang, S., Muramatsu, M., Honjo, T., Nussenzweig, A., and Nussenzweig, M. C. (2004) AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* **118**, 431-438

(6) Xu, Z., Zan, H., Pone, E. J., Mai, T., and Casali, P. (2012) Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat. Rev. Immunol.* **12**, 517-531

(7) de Yébenes, V. G., Belver, L., Pisano, D. G., González, S., Villasante, A., Croce, C., He, L., and Ramiro, A. R. (2008) miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J. Exp. Med.* **205**, 2199-2206

(8) Dorsett, Y., Mcbride, K. M., Jankovic, M., Gazumyan, A., Thai, T.-H., Robbiani, D. F., Di Virgilio, M., Reina-San-Martin, B., Heidkamp, G., Schwickert, T. A., Eisenreich, T., Rajewsky, K., and Nussenzweig, M. C. (2008) MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity* **28**, 630-638

(9) Teng, G., Hakimpour, P., Landgraf, P., Rice, A., Tuschl, T., Casellas, R., and Papavasiliou, F. N. (2008) MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity* **28**, 621-629

(10) Basu, U., Chaudhuri, J., Alpert, C., Dutt, S., Ranganath, S., Li, G., Schrum, J. P., Manis, J. P., and Alt, F. W. (2005) The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature* **438**, 508-511

(11) Mcbride, K. M., Gazumyan, A., Woo, E. M., Barreto, V. M., Robbiani, D. F., Chait, B. T., and Nussenzweig, M. C. (2006) Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 8798-8803

(12) Chatterji, M., Unniraman, S., Mcbride, K. M., and Schatz, D. G. (2007) Role of activation-induced deaminase protein kinase a phosphorylation sites in Ig gene conversion and somatic hypermutation. *J. Immunol.* **179**, 5274-5280

(13) Orthwein, A., Patenaude, A.-M., Affar, E. B., Lamarre, A., Young, J. C., and Di Noia, J. M. (2010) Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp90. *J. Exp. Med.* **207**, 2751-2765

(14) Häsler, J., Rada, C., and Neuberger, M. S. (2011) Cytoplasmic activation-induced cytidine deaminase (AID) exists in stoichiometric complex with translation elongation factor 1α (eEF1A). *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18366-18371

(15) Patenaude, A.-M., Orthwein, A., Hu, Y., Campo, V. A., Kavli, B., Buschiazzo, A., and Di Noia, J. M. (2009) Active nuclear import and cytoplasmic retention of activation-induced deaminase. *Nat. Struct. Mol. Biol.* **16**, 517-527

(16) Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M., and Honjo, T. (2004) Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1975-1980

(17) Mcbride, K. M., Barreto, V., Ramiro, A. R., Stavropoulos, P., and Nussenzweig, M. C. (2004) Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J. Exp. Med.* **199**, 1235-1244

(18) Aoufouchi, S., Faili, A., Zober, C., D'Orlando, O., Weller, S., Weill, J.-C., and Reynaud, C.-A. (2008) Proteasomal degradation restricts the nuclear lifespan of AID. *J. Exp. Med.* **205**, 1357-1368

(19) Uchimura, Y., Barton, L. F., Rada, C., and Neuberger, M. S. (2011) REG-γ associates with and modulates the abundance of nuclear activation-induced deaminase. *J. Exp. Med.* **208**, 2385-2391

(20) Pavri, R., Gazumyan, A., Jankovic, M., Di Virgilio, M., Klein, I., Ansarah-Sobrinho, C., Resch, W., Yamane, A., Reina-San-Martin, B., Barreto, V., Nieland, T. J., Root, D. E., Casellas, R., and Nussenzweig, M. C. (2010) Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* **143**, 122-133

(21) Xu, Z., Fulop, Z., Wu, G., Pone, E. J., Zhang, J., Mai, T., Thomas, L. M., Al-Qahtani, A., White, C. A., Park, S.-R., Steinacker, P., Li, Z., Yates, J., Herron, B., Otto, M., Zan, H., Fu, H., and Casali, P. (2010) 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. *Nat. Struct. Mol. Biol.* **17**, 1124-1135

(22) Nowak, U., Matthews, A. J., Zheng, S., and Chaudhuri, J. (2011) The splicing regulator PTBP2 interacts with the cytidine deaminase AID and promotes binding of AID to switch-region DNA. *Nat. Immunol.* **12**, 160-166

(23) Zheng, S., Vuong, B. Q., Vaidyanathan, B., Lin, J.-Y., Huang, F.-T., and Chaudhuri, J. (2015) Non-coding RNA Generated following Lariat Debranching Mediates Targeting of AID to DNA. *Cell* **161**, 762-773

(24) Arakawa, H., and Buerstedde, J.-M. (2009) Activation-induced cytidine deaminase-mediated hypermutation in the DT40 cell line. *Philos. Trans. R. Soc. Lond, B, Biol. Sci.* **364**, 639-644

(25) Arakawa, H., Saribasak, H., and Buerstedde, J.-M. (2004) Activation-induced cytidine

deaminase initiates immunoglobulin gene conversion and hypermutation by a common intermediate. *Plos Biol.* **2**, E179

(26) Kanehiro, Y., Todo, K., Negishi, M., Fukuoka, J., Gan, W., Hikasa, T., Kaga, Y., Takemoto, M., Magari, M., Li, X., Manley, J. L., Ohmori, H., and Kanayama, N. (2012) Activation-induced cytidine deaminase (AID)-dependent somatic hypermutation requires a splice isoform of the serine/arginine-rich (SR) protein SRSF1. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 1216-1221

(27) Wang, J., Takagaki, Y., and Manley, J. L. (1996) Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability. *Genes Dev.* **10**, 2588-2599

(28) Manley, J. L., and Krainer, A. R. (2010) A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes Dev.* **24**, 1073-1074

(29) Long, J. C., and Cáceres, J. F. (2009) The SR protein family of splicing factors: master regulators of gene expression. *Biochem. J.* **417**, 15-27

(30) Zhong, X.-Y., Wang, P., Han, J., Rosenfeld, M. G., and Fu, X.-D. (2009) SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. *Mol. Cell* **35**, 1-10

(31) Das, S., and Krainer, A. R. (2014) Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Mol. Cancer Res.* **12**, 1195-1204

(32) Cazalla, D., Zhu, J., Manche, L., Huber, E., Krainer, A. R., and Cáceres, J. F. (2002) Nuclear Export and Retention Signals in the RS Domain of SR Proteins. *Mol. Cell. Biol.* **22**, 6871-6882

(33) Wu, J. Y., and Maniatis, T. (1993) Specific Interactions Between Proteins Implicated in Splice-Site Selection and Regulated Alternative Splicing. *Cell* **75**, 1061-1070

(34) Manley, J. L., and Tacke, R. (1996) SR proteins and splicing control. *Genes Dev.* **10**, 1569-1579

(35) Ge, H., Zuo, P., and Manley, J. L. (1991) Primary Structure of the Human Splicing Factor Asf Reveals Similarities with Drosophila Regulators. *Cell* **66**, 373-382

(36) Sun, S., Zhang, Z., Sinha, R., Karni, R., and Krainer, A. R. (2010) SF2/ASF autoregulation involves multiple layers of post-transcriptional and translational control. *Nat. Struct. Mol. Biol.* **17**, 306-312

(37) Zuo, P., and Manley, J. L. (1993) Functional domains of the human splicing factor ASF/SF2. *EMBO J.* **12**, 4727-4737

(38) Conticello, S. G., Ganesh, K., Xue, K., Lu, M., Rada, C., and Neuberger, M. S. (2008) Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNBL1. *Mol. Cell* **31**, 474-484

(39) Basu, U., Meng, F.-L., Keim, C., Grinstein, V., Pefanis, E., Eccleston, J., Zhang, T., Myers, D., Wasserman, C. R., Wesemann, D. R., Januszyk, K., Gregory, R. I., Deng, H., Lima, C. D., and Alt, F. W. (2011) The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* **144**, 353-363

(40) Willmann, K. L., Milosevic, S., Pauklin, S., Schmitz, K.-M., Rangam, G., Simon, M. T., Maslen, S., Skehel, M., Robert, I., Heyer, V., Schiavo, E., Reina-San-Martin, B., and Petersen-Mahrt, S. K. (2012) A role for the RNA pol II-associated PAF complex in AID-induced immune diversification. *J. Exp. Med.* **209**, 2099-2111

(41) Hu, Y., Ericsson, I., Doseth, B., Liabakk, N. B., Krokan, H. E., and Kavli, B. (2014) Activation-induced cytidine deaminase (AID) is localized to subnuclear domains enriched in splicing factors. *Exp. Cell Res.* **322**, 178-192

(42) Rada, C., Jarvis, J. M., and Milstein, C. (2002) AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7003-7008

(43) Barreto, V., Reina-San-Martin, B., Ramiro, A. R., Mcbride, K. M., and Nussenzweig, M.

C. (2003) C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol. Cell* **12**, 501-508

(44) Magari, M., Kanehiro, Y., Todo, K., Ikeda, M., Kanayama, N., and Ohmori, H. (2010) Enhancement of hypermutation frequency in the chicken B cell line DT40 for efficient diversification of the antibody repertoire. *Biochem. Biophys. Res. Commun.* **396**, 353-358

(45) Kanayama, N., Todo, K., Reth, M., and Ohmori, H. (2005) Reversible switching of immunoglobulin hypermutation machinery in a chicken B cell line. *Biochem. Biophys. Res. Commun.* **327**, 70-75

(46) Kanayama, N., Todo, K., Takahashi, S., Magari, M., and Ohmori, H. (2006) Genetic manipulation of an exogenous non-immunoglobulin protein by gene conversion machinery in a chicken B cell line. *Nucleic Acids Res.* **34**, e10

(47) Serrano, P., Aubol, B. E., Keshwani, M. M., Forli, S., Ma, C.-T., Dutta, S. K., Geralt, M., Wüthrich, K., and Adams, J. A. (2016) Directional Phosphorylation and Nuclear Transport of the Splicing Factor SRSF1 Is Regulated by an RNA Recognition Motif. *J. Mol. Biol.* **428**, 2430-2445

(48) Cáceres, J. F., Misteli, T., Screaton, G. R., Spector, D. L., and Krainer, A. R. (1997) Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J. Cell Biol.* **138**, 225-238

(49) Shinkura, R., Ito, S., Begum, N. A., Nagaoka, H., Muramatsu, M., Kinoshita, K., Sakakibara, Y., Hijikata, H., and Honjo, T. (2004) Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat. Immunol.* **5**, 707-712

(50) Ta, V.-T., Nagaoka, H., Catalan, N., Durandy, A., Fischer, A., Imai, K., Nonoyama, S., Tashiro, J., Ikegawa, M., Ito, S., Kinoshita, K., Muramatsu, M., and Honjo, T. (2003) AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat.*

Immunol. **4**, 843-848

(51) Geisberger, R., Rada, C., and Neuberger, M. S. (2009) The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 6736-6741

(52) Ellyard, J. I., Benk, A. S., Taylor, B., Rada, C., and Neuberger, M. S. (2011) The dependence of Ig class-switching on the nuclear export sequence of AID likely reflects interaction with factors additional to Crm1 exportin. *Eur. J. Immunol.* **41**, 485-490

(53) Methot, S. P., Litzler, L. C., Trajtenberg, F., Zahn, A., Robert, F., Pelletier, J., Buschiazzo, A., Magor, B. G., and Di Noia, J. M. (2015) Consecutive interactions with HSP90 and eEF1A underlie a functional maturation and storage pathway of AID in the cytoplasm. *J. Exp. Med.* **212**, 581-596

(54) Nonaka, T., Doi, T., Toyoshima, T., Muramatsu, M., Honjo, T., and Kinoshita, K. (2009) Carboxy-terminal domain of AID required for its mRNA complex formation in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2747-2751

(55) Mondal, S., Begum, N. A., Hu, W., and Honjo, T. (2016) Functional requirements of AID's higher order structures and their interaction with RNA-binding proteins. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E1545-1554

(56) Wu, X., Geraldes, P., Platt, J. L., and Cascalho, M. (2005) The double-edged sword of activation-induced cytidine deaminase. *J. Immunol.* **174**, 934-941

(57) Doi, T., Kato, L., Ito, S., Shinkura, R., Wei, M., Nagaoka, H., Wang, J., and Honjo, T. (2009) The C-terminal region of activation-induced cytidine deaminase is responsible for a recombination function other than DNA cleavage in class switch recombination. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2758-2763

(58) Sabouri, S., Kobayashi, M., Begum, N. A., Xu, J., Hirota, K., and Honjo, T. (2014) C-terminal region of activation-induced cytidine deaminase (AID) is required for efficient

class switch recombination and gene conversion. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2253-2258

(59) Zahn, A., Eranki, A. K., Patenaude, A.-M., Methot, S. P., Fifield, H., Cortizas, E. M., Foster, P., Imai, K., Durandy, A., Larijani, M., Verdun, R. E., and Di Noia, J. M. (2014) Activation induced deaminase C-terminal domain links DNA breaks to end protection and repair during class switch recombination. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E988-997

(60) Ranjit, S., Khair, L., Linehan, E. K., Ucher, A. J., Chakrabarti, M., Schrader, C. E., and Stavnezer, J. (2011) AID recruits UNG and Msh2 to Ig switch regions dependent upon the AID C terminus [corrected]. *J. Immunol.* **187**, 2464-2475

(61) Ucher, A. J., Ranjit, S., Kadungure, T., Linehan, E. K., Khair, L., Xie, E., Limauro, J., Rauch, K. S., Schrader, C. E., and Stavnezer, J. (2014) Mismatch repair proteins and AID activity are required for the dominant negative function of C-terminally deleted AID in class switching. *J. Immunol.* **193**, 1440-1450

(62) Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671-675

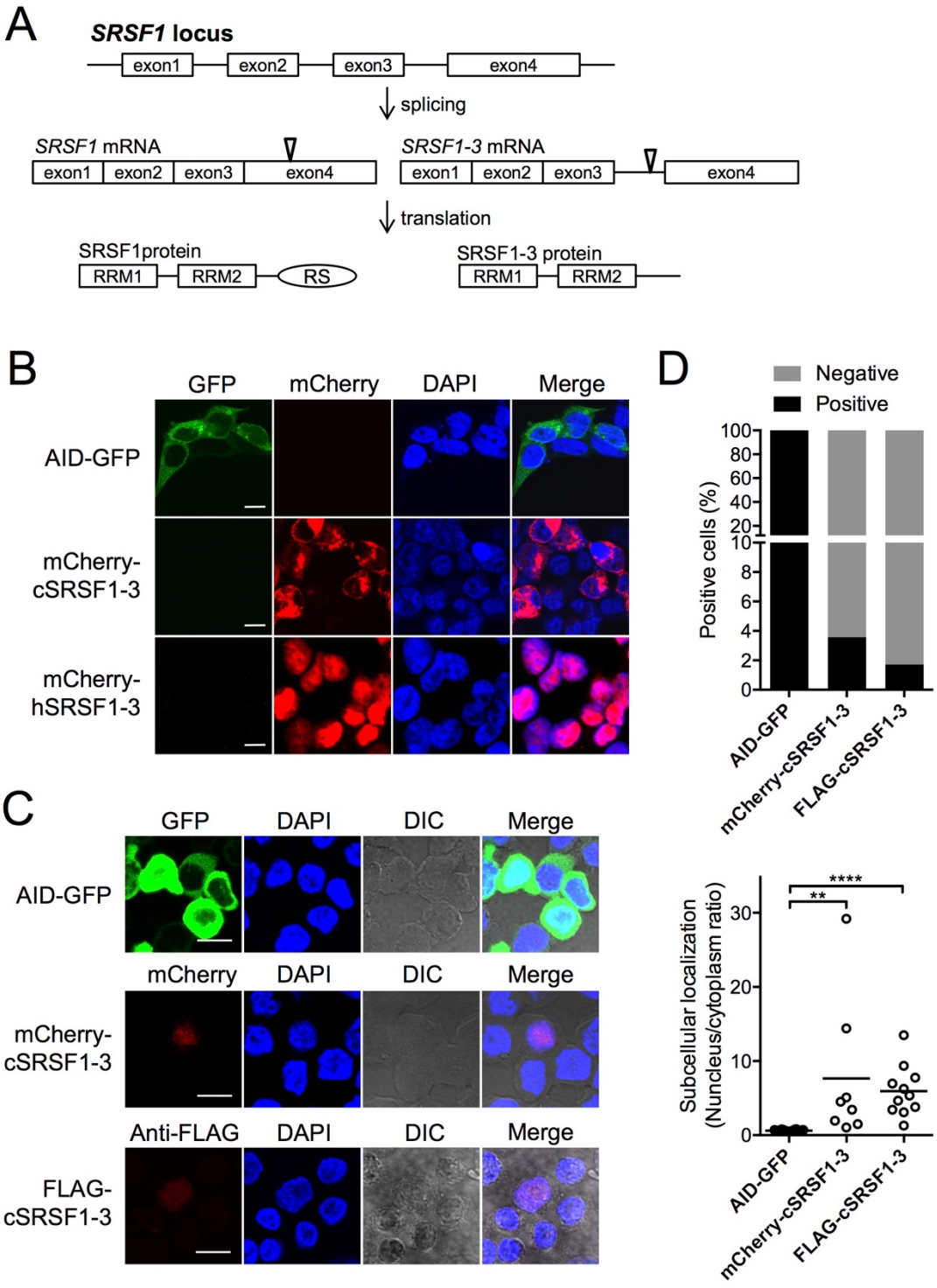

A
SRSF1 locus
exon1
exon2
exon3
exon4
splicing
SRSF1 mRNA
SRSF1-3 mRNA
translation
SRSF1protein
RRM1
RRM2
RS
SRSF1-3 protein
B
GFP
mCherry
DAPI
Merge
AID-GFP
mCherry-cSRSF1-3
mCherry-hSRSF1-3
C
DIC
Anti-FLAG
FLAG-cSRSF1-3
D
Negative
Positive
Positive cells (%)
Subcellular localization (Nuncleus/cytoplasm ratio)
****
**

FIGURE 1. Subcellular localization of AID and SRSF1-3 in 293T cells and DT40 cells.

A, a schematic diagram of the *SRSF1* genomic locus and its products. Arrowheads indicate stop codons on mRNA. RRM, the RNA recognition motif; RS, the RS domain. *B*, in 293T cells, AID is predominantly localized in the cytoplasm and SRSF1-3 is present in the nucleus. Cells were observed by confocal microscopy 36 h after transfection with expression vectors encoding either AID tagged with GFP (AID-GFP), or, chicken or human SRSF1-3 tagged with mCherry (mCherry-cSRSF1-3 or mCherry-hSRSF1-3 respectively). *C*, in DT40 cells, AID and SRSF1-3 are localized in the cytoplasm and the nucleus respectively. DT40-ASF cells stably expressing AID-GFP, mCherry-cSRSF1-3, or FLAG-tagged cSRSF1-3 (FLAG-cSRSF1-3) were used. The scale bar represents 10 μm. DIC, differential interference contrast microscopy. Data are representative of at least two independent experiments. *D*, DT40-ASF cells stably expressing AID-GFP (n = 22), mCherry-cSRSF1-3 (n = 224), or FLAG-cSRSF1-3 (n = 642), which were analyzed in panel C, were evaluated for the frequency of cells expressing detectable AID or SRSF1-3, and the subcellular localization of AID or SRSF1-3 in these cells. Cells with a stronger fluorescence than a background were counted as positive cells, but otherwise counted as negative cells. The proportion of positive cells in total cells were shown in a bar chart (upper panel). The subcellular localization of AID or SRSF1-3 was evaluated by calculating the ratio of nuclear fluorescence over cytoplasmic fluorescence in individual cells, as described in EXPERIMENTAL PROCEDURES (lower panel). The horizontal bar indicates the median of the ratio values for each sample group. **, $p < 0.01$; ****, $p < 0.0001$.

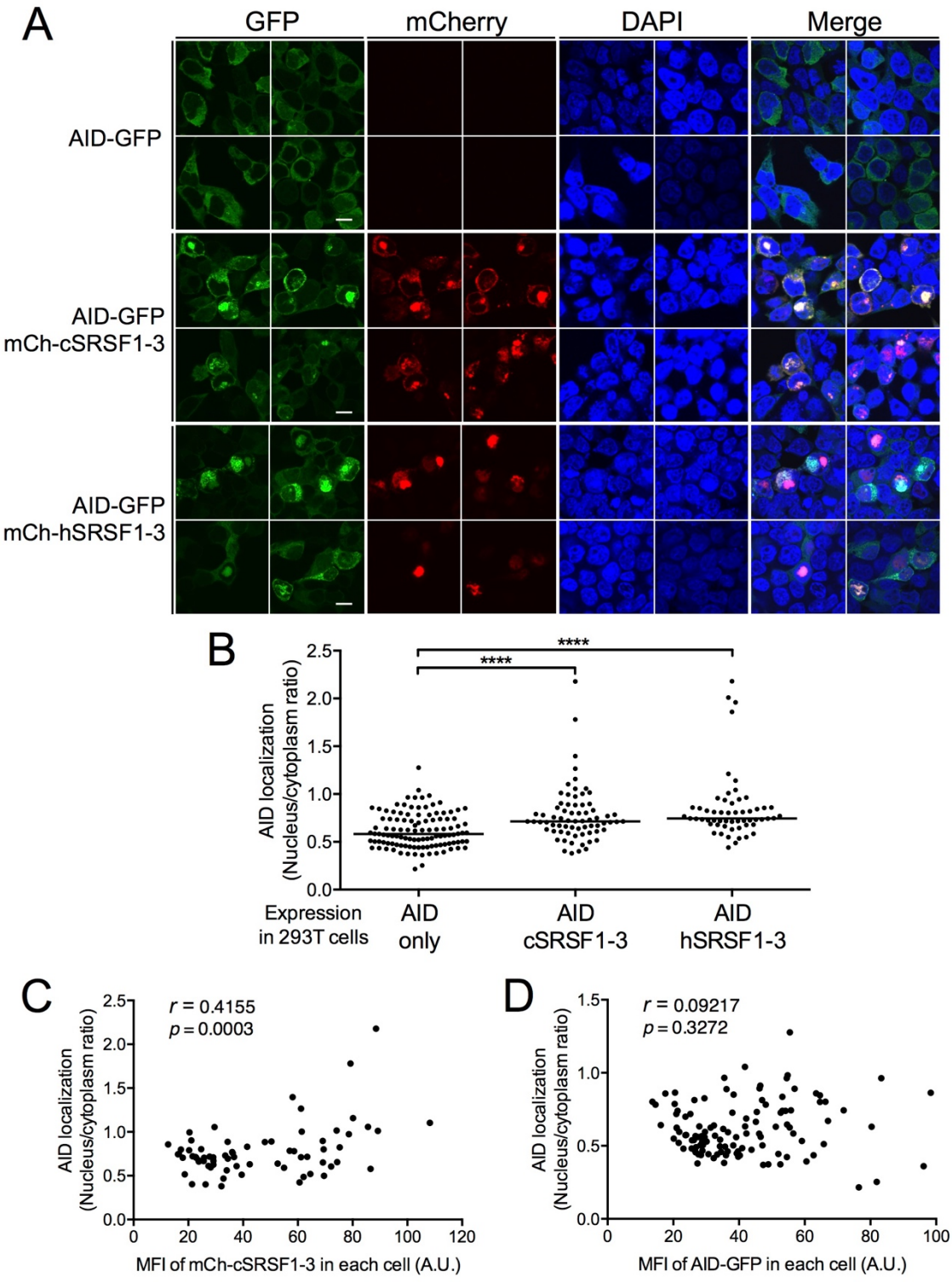

A
GFP
mCherry
DAPI
Merge
AID-GFP
AID-GFP
mCh-cSRSF1-3
AID-GFP
mCh-hSRSF1-3
B
AID localization
(Nucleus/cytoplasm ratio)
2.5
2.0
1.5
1.0
0.5
0.0
****
****
Expression
in 293T cells
AID
only
AID
cSRSF1-3
AID
hSRSF1-3
C
r = 0.4155
p = 0.0003
AID localization
(Nucleus/cytoplasm ratio)
0
20
40
60
80
100
120
MFI of mCh-cSRSF1-3 in each cell (A.U.)
D
r = 0.09217
p = 0.3272
AID localization
(Nucleus/cytoplasm ratio)
1.5
1.0
0.5
0.0
0
20
40
60
80
100
MFI of AID-GFP in each cell (A.U.)

FIGURE 2. Co-expression of AID with SRSF1-3 promotes accumulation of AID in the nucleus.

A, AID accumulates in the nucleus in the presence of overexpressed SRSF1-3. AID-GFP was transiently co-expressed with mCherry-cSRSF1-3 (mCh-cSRSF1-3) or mCherry-hSRSF1-3 (mCh-hSRSF1-3) in 293T cells. The localization of the fusion proteins was observed by confocal microscopy 36 h after transfection with the AID-GFP expression vector alone, or a combination of expression vectors encoding AID-GFP and, mCherry-cSRSF1-3 or mCherry-hSRSF1-3. The scale bar represents 10 μm. *B*, the subcellular localization of AID was evaluated by calculating the ratio of nuclear AID-GFP over cytoplasmic AID-GFP in individual cells, as described in EXPERIMENTAL PROCEDURES. 293T cells expressing AID-GFP alone (n = 115) and co-expressing AID-GFP with mCh-cSRSF1-3 (n = 70) or mCh-hSRSF1-3 (n = 57) were analyzed. The horizontal bar indicates the median of the ratio values for each sample group. ****, $p < 0.0001$. *C*, a correlation analysis for the nuclear localization of AID and the level of SRSF1-3 expression. The correlation analysis was performed using data from 293T cells co-transfected with expression vectors encoding AID-GFP and mCherry-cSRSF1-3. The value for AID localization in each cell shown in panel B was plotted against the mean fluorescent intensity (MFI) of total mCherry-cSRSF1-3 in each cell. The MFI value was calculated from the raw data and expressed as an arbitrary unit (A.U.). *D*, a correlation analysis for the nuclear localization of AID and the level of AID expression. The data from 293T cells transfected with the AID-GFP expression vector alone were used. The value for AID localization in each cell shown in panel B was plotted against the MFI of total AID-GFP in each cell. The MFI value was calculated from the raw data and expressed as an arbitrary unit (A.U.). Data are representative of at least two independent experiments.

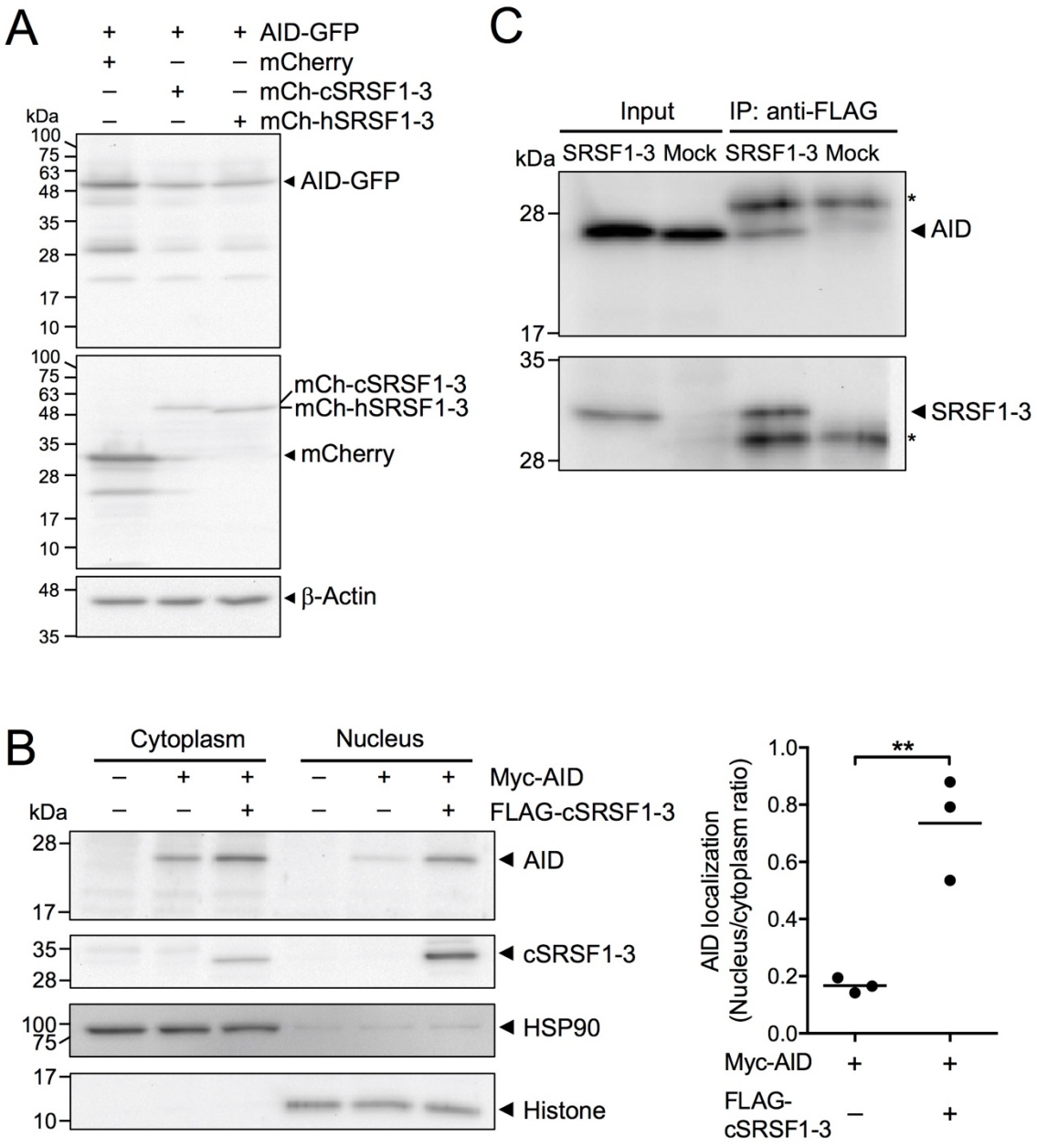

A
+ + + AID-GFP
+ – – mCherry
– + – mCh-cSRSF1-3
– – + mCh-hSRSF1-3
kDa
100
75
63
48
35
28
17
10
AID-GFP
mCh-cSRSF1-3
mCh-hSRSF1-3
mCherry
β-Actin
C
Input
IP: anti-FLAG
kDa SRSF1-3 Mock SRSF1-3 Mock
28
17
35
AID
SRSF1-3
*
B
Cytoplasm
Nucleus
Myc-AID
FLAG-cSRSF1-3
AID
cSRSF1-3
HSP90
Histone
AID localization
(Nucleus/cytoplasm ratio)
1.0
0.8
0.6
0.4
0.2
0.0
**
Myc-AID
FLAG-
cSRSF1-3

FIGURE 3. Biochemical analyses of cells co-expressing AID with SRSF1-3.

A, mCherry-SRSF1-3 fusion proteins do not enhance degradation of AID-GFP fusion protein. AID-GFP was transiently co-expressed with mCherry, mCherry-cSRSF1-3 (mCh-cSRSF1-3), or mCherry-hSRSF1-3 (mCh-hSRSF1-3) in 293T cells. Whole cell lysates were subjected to western blotting using anti-GFP (AID) and anti-mCherry (SRSF1-3) antibodies. ☆-Actin was used as an internal control. *B*, co-expression of AID with SRSF1-3 increases AID in the nuclear fraction. 293T cells transiently transfected with expression vectors encoding Myc-tagged AID (Myc-AID) and FLAG-tagged cSRSF1-3 (FLAG-cSRSF1-3), or with the Myc-AID expression vector and an empty vector, and un-transfected 293T cells were separated into cytoplasmic and nuclear fractions. An aliquot of each fraction was subjected to western blotting using anti-AID and anti-FLAG (SRSF1-3) mAbs (left panel). Hsp90 and acetylhistone H4 were used as internal controls for the cytoplasmic and nuclear fractions, respectively. The signal intensity of AID was quantified and the nuclear localization of AID was expressed as the ratio of nuclear AID over cytoplasmic AID (right panel). The horizontal bar indicates the mean of the ratio values for three independent experiments. **, $p < 0.01$. *C*, SRSF1-3 forms a protein complex with AID. Cell lysates were prepared from 293T cells co-transfected with expression vectors encoding Myc-AID and FLAG-cSRSF1-3 (SRSF1-3), or with the Myc-AID expression vector and an empty vector (Mock). Protein complexes containing FLAG-SRSF1-3 were immunoprecipitated using an anti-FLAG mAb, and analyzed by western blotting using anti-Myc (AID) or anti-FLAG (SRSF1-3) mAbs. Asterisks indicate the L chain of the anti-FLAG mAb used for immunoprecipitation. Data are representative of three independent experiments.

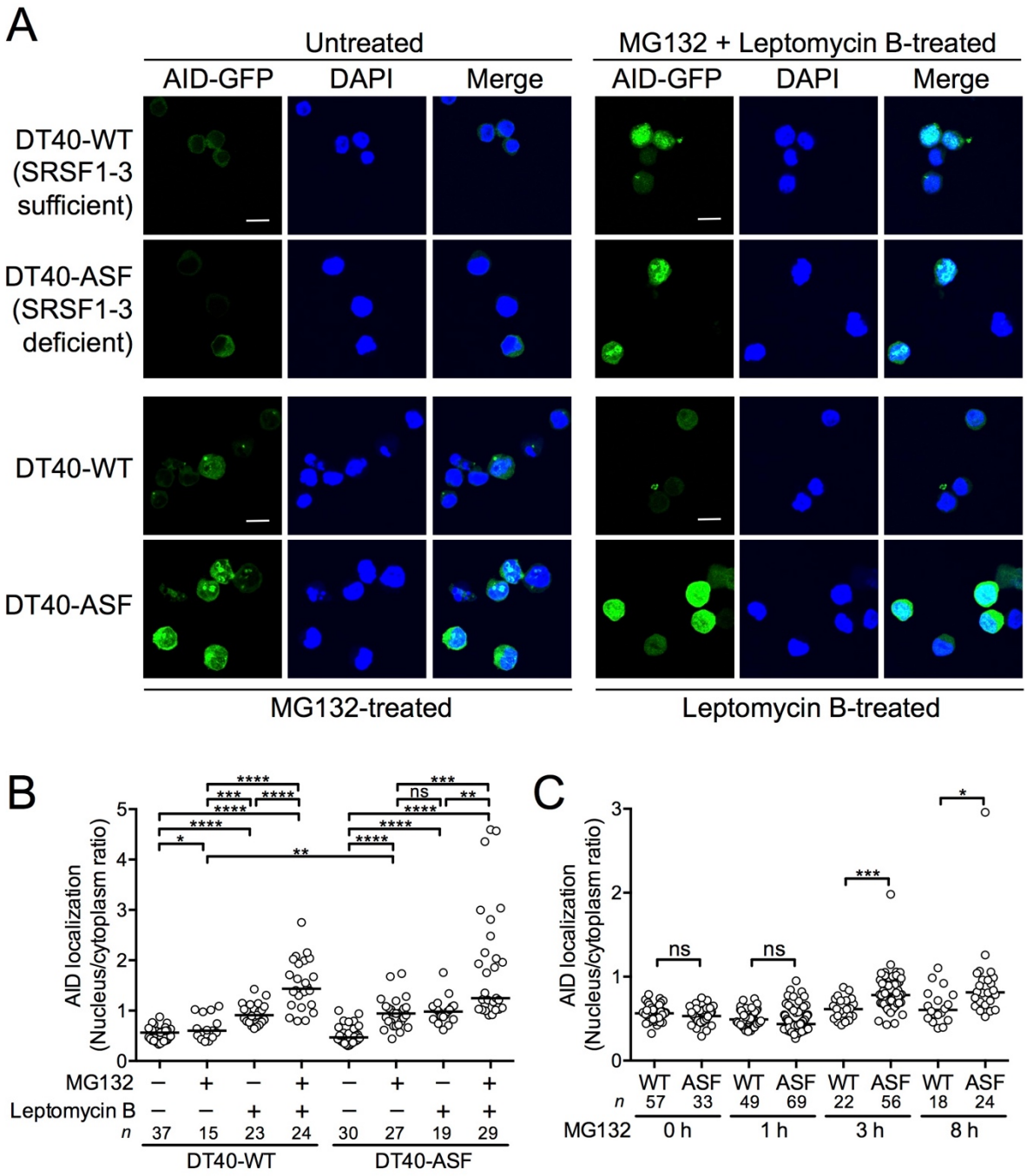

A
Untreated
MG132 + Leptomycin B-treated
AID-GFP
DAPI
Merge
DT40-WT (SRSF1-3 sufficient)
DT40-ASF (SRSF1-3 deficient)
DT40-WT
DT40-ASF
MG132-treated
Leptomycin B-treated
B
AID localization (Nucleus/cytoplasm ratio)
MG132
Leptomycin B
n
37 15 23 24 30 27 19 29
DT40-WT
DT40-ASF
C
AID localization (Nucleus/cytoplasm ratio)
WT ASF WT ASF WT ASF WT ASF
n 57 33 49 69 22 56 18 24
MG132 0 h 1 h 3 h 8 h
ns
*
**
***
****

FIGURE 4. SRSF1-3 is dispensable for nuclear import of AID.

A, AID accumulated in the nucleus by inhibiting protein degradation and/or nuclear export either in the presence or absence of SRSF1-3 in DT40 cells. SRSF1-3-sufficient (DT40-WT) and SRSF1-3-deficient (DT40-ASF) DT40 cells stably expressing AID-GFP were treated with MG132 (6 h) and/or leptomycin B (2 h) to inhibit proteasomal degradation and nuclear export, respectively, and then observed by confocal microscopy. The scale bar represents 10 μm. Data are representative of two independent experiments. *B*, the subcellular localization of AID in DT40 cells untreated or treated with MG132 and/or leptomycin B (panel A) was evaluated by calculating the ratio of nuclear AID-GFP over cytoplasmic AID-GFP in individual cells, as described in EXPERIMENTAL PROCEDURES. *C*, the time course of the subcellular localization of AID in DT40 cells untreated or treated with MG132 and/or leptomycin B. SRSF1-3-sufficient (DT40-WT, WT) and SRSF1-3-deficient (DT40-ASF, ASF) cells stably expressing AID-GFP were treated with MG132 for indicated time periods, and then observed by confocal microscopy. The subcellular localization of AID in each cell was evaluated as described above. The horizontal bar indicates the median of the ratio values for each sample group. The statistical significance between two groups is indicated, and there was no significant difference between the other combinations (*B* and *C*). *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$; ns, not significant.

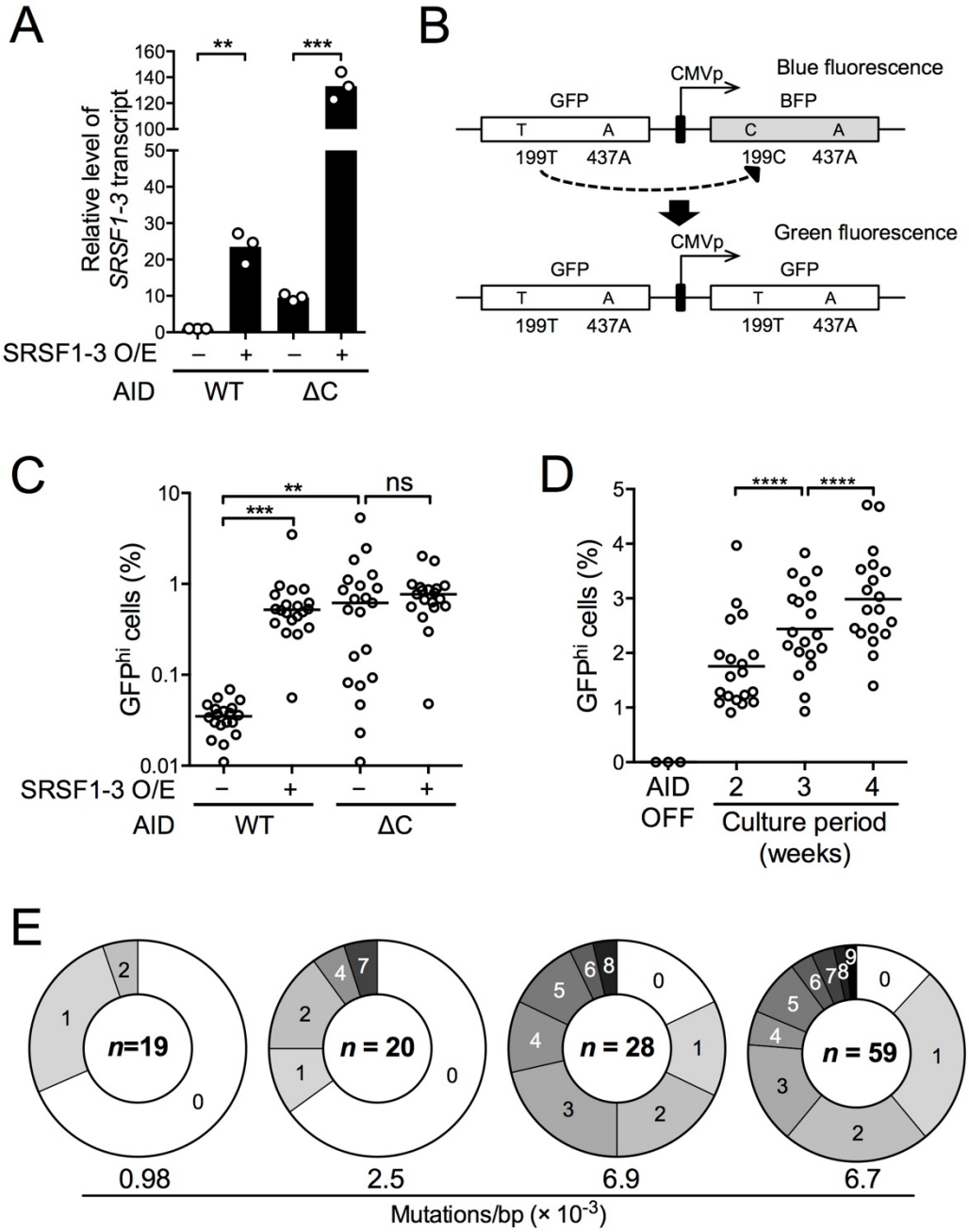
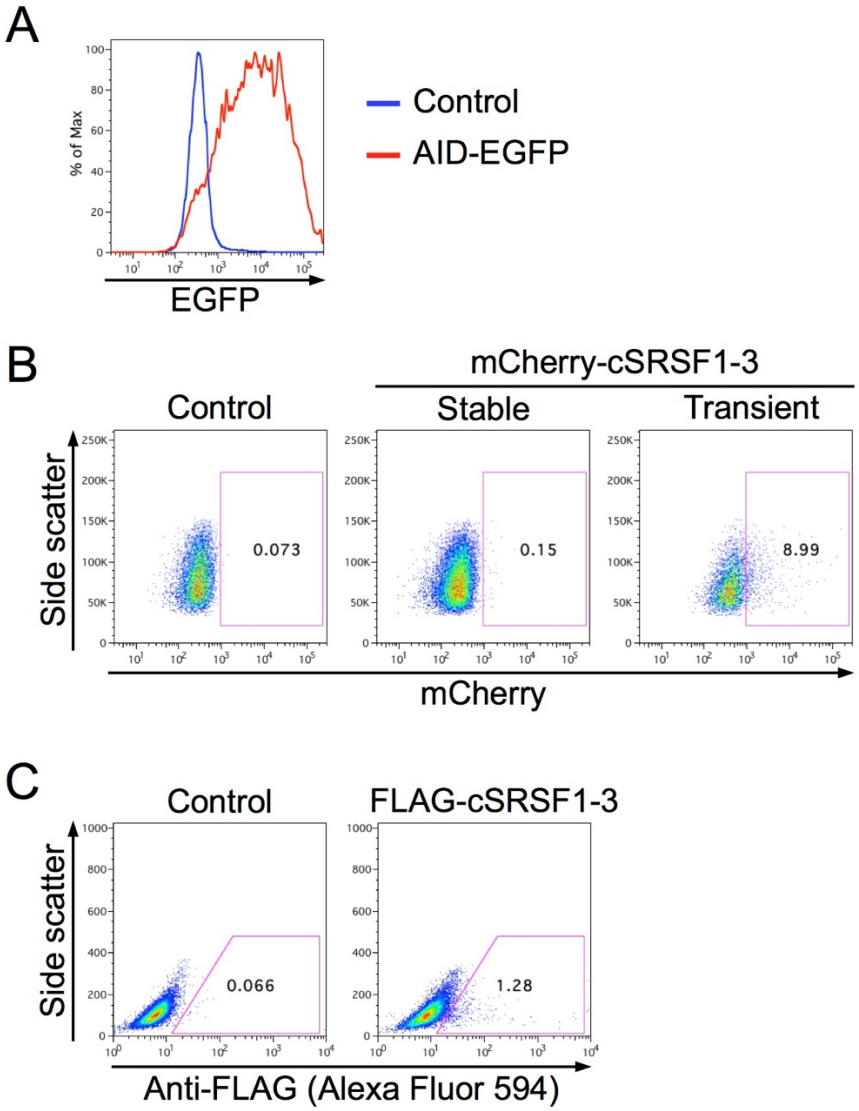

A
Relative level of SRSF1-3 transcript
160
140
120
100
50
40
30
20
10
0
**
***
SRSF1-3 O/E
–
+
–
+
AID
WT
ΔC
B
CMVp
Blue fluorescence
GFP
BFP
T
A
C
A
199T
437A
199C
437A
Green fluorescence
CMVp
GFP
GFP
T
A
T
A
199T
437A
199T
437A
C
GFPhi cells (%)
10
1
0.1
0.01
**
ns
***
SRSF1-3 O/E
–
+
–
+
AID
WT
ΔC
D
GFPhi cells (%)
5
4
3
2
1
0
****
****
AID OFF
2
3
4
Culture period (weeks)
E
n=19
2
1
0
0.98
n = 20
4
7
2
1
0
2.5
n = 28
6
8
5
0
4
1
3
2
6.9
n = 59
6
7
8
9
5
0
4
3
1
2
6.7
Mutations/bp (× 10-3)

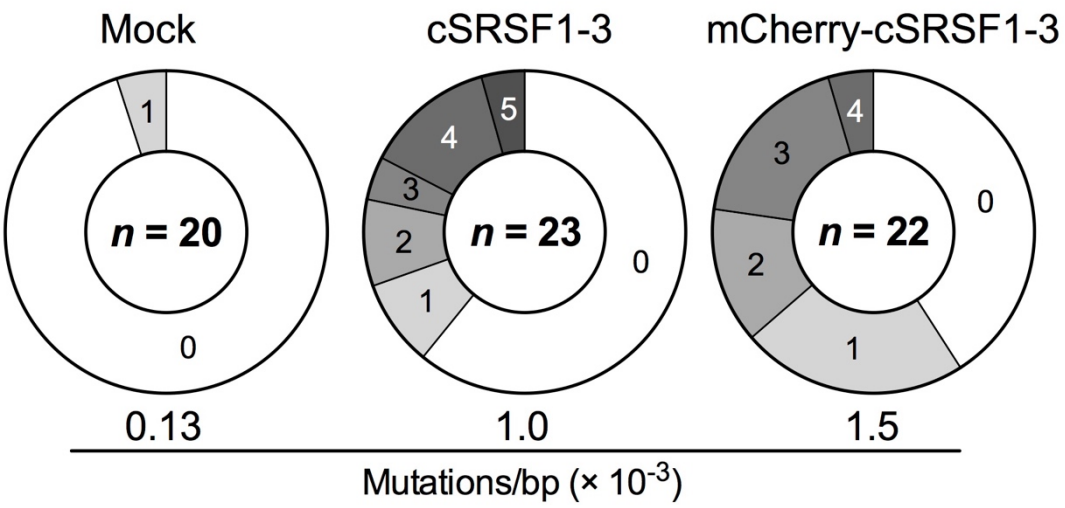
FIGURE 5. A role for SRSF1-3 in *IgV* diversification depends on the C-terminal region of AID.

A, the transcript level of SRSF1-3 in DT40 cells stably overexpressing SRSF1-3. The SRSF1-3 expression vector was introduced into DT40 cells expressing the wild-type AID-expressing (AID WT) or a C-terminally-deleted AID (AIDΔC). The level of *SRSF1-3* transcript was analyzed by qRT-PCR. Values are means ± SD (n = 3). **, $p < 0.005$; ***, $p < 0.0005$. *B*, G/B construct is an artificial substrate that can easily be used to efficiently monitor the frequency of GCV in DT40 cells (46). A single nucleotide change in the BFP gene, which is templated from the GFP gene, results in a fluorescence shift from blue to green. *C*, overexpression of SRSF1-3 does not further enhance GCV activity in AIDΔC-expressing DT40 cells. BFP^+ (GCV-un-induced) and GFP^{lo} (AID-expressing) single cells (n = 19 or 20) were sorted from DT40 cells expressing the wild-type AID or AIDΔC, with or without overexpression of SRSF1-3, and cultured for 2 weeks. GFP^{hi} (GCV-induced) cells were counted by flow cytometry. The horizontal bar indicates the median of GFP^{hi} cell proportions for each sample group. **, $p < 0.005$; ***, $p < 0.0005$; ns, not significant. Data are representative of three independent experiments. *D*, the time course of GCV occurrence in AIDΔC-expressing DT40 cells. BFP^+ (GCV-un-induced) and GFP^{lo} (AID-expressing) single cells (n = 19) were sorted from DT40 cells expressing AIDΔC without overexpression of SRSF1-3, and cultured for 2, 3, and 4 weeks. GFP^{hi} (GCV-induced) cells were counted by flow cytometry. DT40 cells without AID expression were used as a negative control. *E*, overexpression of SRSF1-3 does not further enhance hypermutation on the *IgVH* gene in AIDΔC-expressing DT40 cells. The *IgVH* genes of DT40 cells expressing the wild-type AID or AIDΔC, with or without overexpression of SRSF1-3, were cloned and sequenced after 3 weeks of culture. Segment sizes in pie charts show the proportion of sequences with 0-9 mutations.


A
% of Max
EGFP
Control
AID-EGFP
B
mCherry-cSRSF1-3
Control
Stable
Transient
Side scatter
mCherry
0.073
0.15
8.99
C
Control
FLAG-cSRSF1-3
Side scatter
Anti-FLAG (Alexa Fluor 594)
0.066
1.28

Supplemental FIGURE S1. Flow cytometric analysis of the expression of AID-EGFP, mCherry-cSRSF1-3, or FALG-cSRSF1-3 in stably transfected DT40-ASF cells.

A, the expression of AID-EGFP in DT40-ASF cells stably transfected with the AID-EGFP expression vector. Untransfected DT40-ASF cells were used as a control. B, the expression of mCherry-cSRSF1-3 in DT40-ASF cells stably or transiently transfected with the mCherry-cSRSF1-3 expression vector. Cells 24 h after transfection were analyzed by flow cytometry. *C*, the expression of FLAG-cSRSF1-3 in DT40-ASF cells stably transfected with the FLAG-cSRSF1-3 expression vector. Control and transfected cells were fixed with paraformaldehyde, permeabilized with nonidet P-40, and stained with a mouse anti-FLAG mAb (M2) followed by Alexa Fluor 594-conjugated rabbit anti-mouse IgG antibody.



Supplemental FIGURE S2. mCherry-cSRSF1-3 is competent to induce *IgV* hypermutation comparably to the authentic SRSF1-3.

The *IgVH* genes of DT40-ASF cells bearing a mock vector, or the expression vectors encoding cSRSF1-3 or mCherry-cSRSF1-3, were cloned and sequenced after 30 days of culture. Segment sizes in pie charts show the proportion of sequences with 0-9 mutations.

List of Pubulication

1. SRSF1-3 Contributes to Diversification of the Immunoglobulin Variable Region Gene by Promoting Accumulation of AID in the Nucleus

   Yuka Kawaguchi, Hiroaki Nariki, Naoko Kawamoto, Yuichi Kanehiro, Satoshi Miyazaki, Mari Suzuki, Masaki Magari, Hiroshi Tokumitsu, Naoki Kanayama

   Biochemical and Biophysical Research Communications, in press.

   DOI: http://dx.doi.org/10.1016/j.bbrc.2017.02.097

Acknowledgments

These studies collected here have been carried out under the direction of Associate Professor Naoki Kanayama in Laboratory of Applied Cell Biology, Department of Biotechnology, Graduate School of Natural Science and Technology, Okayama University.

The author greatly appreciates Dr. Naoki Kanayama of Okayama University for kind support throughout my research.

The author also appreciates Prof. Hiroshi Tokumitsu of Okayama University for kind advice.

The author also appreciates Dr. Masaki Magari of Okayama University for helpful suggestions.

The author also appreciates Prof. Hitoshi Ohmori of Okayama University for kind support.

The author also appreciates Dr. Yuichi Kanehiro of Okayama University for helpful advice

Acknowledgements are made to Mr. Hiroaki Nariki, Ms. Naoko Kawamoto, Mr. Kazuki Yokoyama, Dr. Ayano Satoh, Mr. Satoshi Miyazaki, Ms. Machi Yoneda, Mr. Yuma Matsuyama, Ms. Mari Suzuki for their kind technical assistance throughout this study.

Finally, thanks are due to the member of the Laboratory of Applied Cell Biology, Department of Biotechnology, Graduate School of Natural Science and Technology, Okayama University, for their helpful discussion and kind supports.