Active nuclear import and cytoplasmic retention of Activation Induced Deaminase

Anne-Marie Patenaude[1], Alexandre Orthwein[1,2], Yi Hu[1,3], Vanina A Campo[1], Bodil Kavli[3], Alejandro Buschiazzo[4,5] and Javier M Di Noia[1,2,6]

[1]Institut de Recherches Cliniques de Montréal, 110 Av des Pins Ouest, Montréal, QC, H2W 1R7, Canada.

[2]Department of Microbiology and Immunology, Université de Montréal, Montréal, QC, Canada.

[3]Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Erling Skjalgssons gt. 1, DMF, 7006 Trondheim, Norway.

[4]Institut Pasteur de Montevideo, Unidad de Cristalografia de Proteinas, Montevideo, Uruguay. Mataojo 2020, Montevideo 11400, Uruguay.

[5]Institut Pasteur, Département de Biologie Structurale & Chimie, Paris, France.

[6]Department of Medicine, University of Montreal, Montreal, QC, Canada.

Correspondence should be addressed to J.M.D email: javier.di.noia@ircm.qc.ca Tel: 1-514-987-5642, Fax: 1-514-987-5645.

Abstract

The enzyme Activation Induced Deaminase (AID) triggers antibody diversification in B-cells by catalyzing deamination and consequently mutation of immunoglobulin genes. To minimize off-target deamination, AID is restrained by several regulatory mechanisms including nuclear exclusion, thought to be mediated exclusively by active nuclear export. Here we identify two other mechanisms involved in controlling AID subcellular localization. AID is unable to passively diffuse into the nucleus, despite its small size, its nuclear entry requiring active import mediated by a conformational nuclear localization sequence (NLS). We also identify a determinant for AID cytoplasmic retention in its C-terminus, which hampers diffusion to the nucleus, competes with nuclear import and is critical for maintaining the predominantly cytoplasmic localization of AID in steady-state conditions. Blocking nuclear import alters the balance between these processes in favor of cytoplasmic retention, resulting in reduced isotype class switching.

INTRODUCTION

Activation Induced Deaminase (AID) is the enzyme responsible for the diversification of rearranged antibody genes in activated B cells[1,2]. AID acts as a mutator by deaminating deoxycytidine in single stranded DNA thereby changing the base, cytosine, into a uracil (reviewed in[3]). Through further processing by DNA repair enzymes that recognize uracil in DNA, this single biochemical activity triggers different genetic modifications that are critical for a proper antibody response (reviewed in[3,4]). Thus, AID initiates somatic hypermutation (SHM) and Ig gene conversion at the immunoglobulin (Ig) variable regions allowing for affinity maturation of the antibody response[1,2]. AID also initiates class switch recombination (CSR), which exchanges the exons encoding the Fc region of the antibody from the default IgM to another isotype[1,2]. All the known components in these pathways except for AID are ubiquitous DNA repair enzymes (reviewed in[3-5]). Indeed, AID is able to trigger SHM and CSR in non-B cell models[6,7] and since such mutagenic and recombinogenic enzyme is potentially dangerous, it needs to be tightly regulated. This is highlighted by the cancer predisposition phenotype observed in transgenic mice overexpressing AID[7], by the finding that ectopic SHM can occur in proto-oncogenes and tumor suppressor genes[8-10] and by the involvement of AID in oncogenic chromosomal translocations[11,12]. AID is normally induced in germinal center B cells[13] but in order to ensure that the genetic modifications it can cause are largely restricted to the Ig loci, there are multiple points of post-transcriptional regulation such as regulation of mRNA stability and translation[14-16], subcellular localization[17,18], protein stability[19] and modification by phosphorylation[20-22].

The subcellular localization of AID is especially interesting. The initially puzzling observation that an AID-GFP fusion was exclusively cytoplasmic[23] could be later explained by the fact that AID is a nuclear-cytoplasmic shuttling protein. The mechanism by which AID is transported out of the nucleus is known; a Leucine-rich nuclear export signal (NES) at the C-terminus of AID is recognized by the exportin CRM1[17,18,24]. However, the mechanism by which AID enters the nucleus has not been studied in any detail and the observations available are inconsistent. While one report proposes a classical bipartite nuclear localization signal (NLS) at the N-terminus of AID[17], others failed to find evidence for such a signal, suggesting that AID may simply diffuse into the nucleus[18,24]. We have investigated the mechanism by which human AID enters the nucleus to address this controversy. We find that, despite its small size, AID is actually unable to passively diffuse into the nucleus, requiring instead active import. In addition, we describe a novel cytoplasmic retention mechanism for AID, which contributes to the observed nuclear exclusion in steady state, thus restraining its function.

RESULTS

AID is actively imported into the nucleus

The expression of AID-GFP in several cell types resulted in exclusive cytoplasmic localization in steady state, with this changing to nuclear localization after inhibition of CRM1-mediated nuclear export by leptomycin B, as previously reported[17,18,24]. We indeed observed that AID-GFP accumulated into the nucleus after leptomycin B treatment in the majority of transiently transfected 293T (80% after 4 h, $n = 206$) and HeLa (96% after 1 h, $n = 53$) cells and in stably expressing Ramos B cells (93% after 1 h, $n = 91$). Moreover, AID-GFP accumulated in the nucleus with different kinetics depending on the cell line (Fig. 1 a). These observations are not consistent with the previous proposal that AID subcellular localization is determined exclusively by active nuclear export and nuclear entry occurring by passive diffusion[18,24]. In that case AID-GFP would be expected to follow the mass action law and reach a more homogeneous distribution between nucleus and cytoplasm after leptomycin B treatment. Rather, additional mechanisms seem to control AID localization.

Since active translocation across nuclear pores requires energy[25,26], nuclear proteins that are not retained should diffuse out of the nucleus when ATP is depleted[27]. We depleted ATP from HeLa cells by glucose deprivation and assessed the localization of AID L198S-GFP, a constitutively nuclear AID variant carrying a point mutation inactivating the NES[18], and NLS^{SV40}-GFP. The latter was used as a positive control since it is small enough to diffuse through the nuclear pores and its nuclear localization is mediated only by the NLS. ATP depletion led to redistribution of NLS^{SV40}-GFP throughout the cell in >90% of the cells and ~40% of the cells transfected with AID L198S-GFP (Fig. 1 b). Importantly, this redistribution was reversible as both proteins accumulated back into the nucleus upon transferring the cells to complete medium (Fig. 1 b). The differential sensitivity of AID L198S-GFP to this treatment compared to NLS^{SV40}-GFP may suggest nuclear retention of the former, which could be due to a larger protein size or binding to DNA[28] or to nuclear factors[29], as well as a previously proposed retention by DNA damage[24]. We did not pursue this observation further. Nevertheless, the fact that energy was needed to maintain AID nuclear localization in a substantial proportion of cells supports the existence of active nuclear import.

The ~50 kDa AID-GFP is at the upper limit of the nuclear pore cut-off[29,30]. We confirmed that AID could mediate active nuclear import by increasing the size of the fusion protein so as to preclude diffusion. Since protein shape may strongly influence the ability to diffuse[30], we used APOBEC2 (A2) as control. Given the homology with A2 and APOBEC3G[31-33], the predicted three-dimensional

4

structure of AID safely allows us to postulate that the monomers of A2 (25.7 kDa) and AID (23.9 kDa) will have a similar general folding, and therefore shape (Supplementary Fig. 1). Several controls confirmed that A2-GFP has no information for nuclear import or export and therefore its throughout the cell localization is achieved by passive diffusion (Fig. 1 c and Supplementary Fig. 2). Thus, increasing the size of the A2-GFP to ~75 kDa by adding β-Lactamase (β-Lac) or to ~175 kDa using β-Galactosidase (β̄-Gal) resulted in exclusive cytoplasmic localization, even after leptomycin B treatment. Instead, AID-β-Lac-GFP accumulated in the nucleus upon leptomycin B treatment and the export deficient mutants AID L198S-β-Lac-GFP and AID 181-β-Lac-GFP (that lacks the last 17-residues of AID) were constitutively nuclear (Fig. 1 c). Moreover, AID 181 was also able to change the subcellular localization of β-Gal-GFP (~150 kDa) from cytoplasmic to nuclear, just as the *bona fide* nuclear protein UNG2[34] did (Fig. 1 d). We conclude that AID nuclear import is active.

Most of AID protein is necessary to mediate nuclear import

The existence of an NLS in AID is controversial[17,18,24]. We systematically investigated whether AID contained a sequence that could be sufficient to act as an NLS. To this end, we generated a collection of nested C-terminal deletions of AID fused to the N-terminus of GFP (Fig. 2 a). Most of these fusion proteins were distributed throughout cells. Only AID 181 and 187 fully transported GFP into the nucleus of 293T, Hela and Ramos cells (Fig. 2 a, b and c and Supplementary Fig. 3). To further examine whether the N-terminal domain of AID contained an NLS without being obscured by the effects of diffusion, we increased the size of the fusion proteins by including β-Gal. Both, AID 40-β-Gal-GFP and AID 54-β-Gal-GFP showed only weak nuclear signals while the bipartite NLS from nucleoplasmin and AID 181 efficiently transported β-Gal-GFP into the nucleus of 293T and Hela cells (Fig. 2 d and not shown).

To define the regions of AID that are relevant for nuclear import by a different approach, we used a set of five chimeric proteins (AID-A2 #1 to #5), in which different regions of AID were replaced by the homologous A2 regions[35], with a C-terminal GFP tag (Fig. 2 e and Supplementary Fig. 4). AID-A2 #1, in which amino acids 19–57 of AID were replaced with residues 60–96 of A2, remained exclusively cytoplasmic even after nuclear export establishing this N-terminal domain as critical for nuclear import. AID-A2 #2 localized throughout the cells after leptomycin B treatment with AID-A2 #3 and #4 showing similar but milder defects for nuclear accumulation in HeLa cells (Fig. 2 e), which were more evident in 293T cells (Supplementary Fig. 4). Since AID-A2 #2, #3 and #4 were catalytically inactive (Supplementary Fig. 4), a structural defect could explain the effect seen on import. AID-A2 #5 had the

5

C-termini interchanged; replacing AID positions 154–198 (thus deleting the NES) with the last 31 amino acids of A2 (which has no detectable NES). Despite this exchange, AID-A2 #5 was not nuclear (as it would be expected if it had a linear NLS) but distributed throughout the cell. AID-A2 #5-β-Gal-GFP was cytoplasmic (Supplementary Fig. 4) confirming loss of active import. Importantly, AID-A2 #5 was active (Supplementary Fig. 4) ruling out a major folding alteration as a cause to prevent import. Rather, it is consistent with the results from the C-terminal truncations indicating that a minimal length of AID (>160 residues) is required to allow efficient active nuclear import and suggesting the need for structural integrity.

AID has a conformational positively charged NLS

Since AID is likely to oligomerize (see Discussion), we explored the possible influence of AID quaternary structure on nuclear import. AID dimerization through the β2 strand has been proposed based on the structure of A2[33]. We tested the effect of perturbing the predicted AID β2, residues 40-53 in our model, on oligomerization by introducing one (F46A), two (F46A/Y48A)[33] or four (F46A/Y48A/R50G/N51A, named AID FYRN) mutations. The ability of each of these mutants to interact with wt AID was monitored by comparing their efficiency in coimmunoprecipitating with AID-Flag. While AID F46A-GFP and AID F46A/Y48A-GFP coimmunoprecipitated similarly to AID-GFP in this assay, AID FYRN-GFP did so much less efficiently or not at all depending on the experiment (Fig. 3 a and not shown). When AID FYRN-Flag was used as bait AID FYRN-GFP or AID-GFP failed to coimmunoprecipitate (Fig. 3 a). Failure to oligomerize correlated well with defective nuclear import. While AID F46A-GFP and AID F46A/Y48A-GFP exhibited normal shuttling, AID FYRN-GFP was distributed throughout cells after leptomycin B treatment (Fig. 3 b and Supplementary Fig. 5), suggesting defective nuclear import. Indeed, AID FYRN with a truncation of the last 17 amino acids was unable to transport β-Gal into the nucleus while the equivalent unmutated AID 181-β-Gal-GFP did (Fig. 3 b). However, AID FYRN showed a shorter half-life than AID and it was inactive (Supplementary Fig. 5) so a structural defect cannot be excluded as causing the import deficiency. In any case, the failure of AID FYRN to accumulate in the nucleus could not be explained by the diminished stability of the protein since it was still excluded from the nucleus in steady state (*i.e.* export was active) and could be complemented for nuclear accumulation by the addition of NLS[SV40] (Supplementary Fig. 5).

While investigating the effect of mutations in the putative oligomerization interfaces of AID[33] we found three other Arg residues that were important for nuclear import in addition to Arg50 in FYRN.

Mutations R19A, R24W and R112D caused import defects while mutations at Tyr114/Phe115, Glu117/Asp118 did not (Supplementary Fig. 5). Therefore, we examined in more detail the positively charged N-terminal domain of AID by using chimeras in which this domain was substituted in segments of 3–5 residues with the corresponding A2 positions[35]. Only replacements 19–22, 34–36 or 50–54, all involving basic residues, caused mislocalization of AID after nuclear export inhibition (Fig. 3 c). The combination of two of these replacements in AID-A2 19–22/34–36 and AID-A2 34–36/50–54, showed a more drastic effect than any single one resulting in persistent cytoplasmic localization after leptomycin B treatment (Fig. 3 d). Confirming the importance of these determinants for import, untagged AID-A2 #1 or AID-A2 34–36 remained largely cytoplasmic after leptomycin B (Fig. 3 e). Here again the effect of replacing residues 34–36 was less drastic than replacing 19–57 in #1. Thus, at least 4 determinants involving basic residues collaborate in mediating efficient nuclear import of AID.

The fact that AID requires several non-consecutive determinants, as well as a substantial length of the protein, for efficient nuclear import strongly suggests a conformational NLS. In fact, when the residues affecting AID import were displayed on the 3D model it became apparent that they were all in close proximity within the same surface area except for residues 34-36, which would however be included in the same surface in a putative AID dimer (Fig. 4 a). A number of the basic residues are exposed in this model suggesting a positively charged, classical NLS (Fig. 4 b). This kind of NLS is recognized by members of the importin-α family of adaptors[29,36]. Indeed, AID-Flag was pulled down from extracts of Ramos cells by GST-importins-α1, α3 and α5 (Fig. 4 c). We obtained identical results with AID-GFP (not shown). In agreement with our localization results AID-GFP and the constitutively nuclear AID 181-GFP were pulled down by GST-importin-α3 while A2-GFP control or the import-deficient AID mutants AID FYRN and AID with an N-terminal truncation were not (Fig. 4 d). Altogether, our results are consistent with a conformational classical NLS determined by the protein folding and, possibly, oligomerization.

Hindered diffusion of AID from the cytoplasm

The observation that several import-deficient AID-GFP variants (AID-A2 19–22, 19–22/34–36 and 34–36/50–54 and AID R19A) and even untagged AID-A2 #1 remained mainly cytoplasmic after nuclear export inhibition, suggested that AID was unable to freely diffuse into the nucleus. We confirmed that AID-A2 #1-GFP remained exclusively cytoplasmic after leptomycin B treatment also in B cells while the similarly sized and shaped A2-GFP was localized throughout the cells (Fig. 5 a). The finding that N-terminal tagging of unmutated AID apparently prevented its nuclear translocation reinforced the possibility of cytoplasmic retention. Thus, GFP-AID and even Flag-AID (~25kDa) failed to enter into

7

the nucleus of B and non-B cells after leptomycin B treatment (Fig. 5 b and Supplementary Fig. 6). N-terminal fusions do not seem to affect AID oligomerization[35] (Supplementary Fig. 6) or catalytic activity[28,33,35]. Rather, the position of the N-terminus respect to the import determinants suggests that they might be masking the NLS (Fig. 4 a). Regardless of the explanation, the possibility that the persistent cytoplasmic localization of GFP-AID or AID-A2 #1-GFP simply reflects reduced active nuclear import is very unlikely since both remained cytoplasmic up to 6 h post leptomycin B treatment (Fig. 5 c and Supplementary Fig. 6). By this time, they should have been able to at least passively diffuse into the nucleus given their size and as demonstrated by the fact that of AID L198S-GFP diffused out of the nucleus within 3 h in the energy depletion experiments. Finally, we took advantage of the general inhibition of importin-mediated nuclear import produced by oxidative stress[37,38] to further test the cytoplasmic retention of unmutated AID after leptomycin B treatment. Indeed, AID-GFP failed to enter the nucleus in response to leptomycin B in Hela cells that had been pretreated with hydrogen peroxide although the distribution of A2-GFP was not affected (Fig. 5 d). Even untagged AID, which is well below the nuclear pore cut-off, remained cytoplasmic in these conditions (Fig. 5 e). Altogether, our results strongly suggest that the diffusion of AID from the cytoplasm is hindered, indicating the existence of a retention mechanism.

A C-terminal cytoplasmic retention determinant

To identify regions of AID mediating cytoplasmic retention we used again AID-A2 chimeras but N-terminally GFP tagged so as to block import. Replacing any region of AID contributing to mediate cytoplasmic retention should, at least by passive diffusion, allow some nuclear localization after leptomycin B treatment. Only GFP-AID-A2 #5 was able to access the nucleus (Fig. 6 a) suggesting that AID residues 158 to 198 were necessary for cytoplasmic retention. Indeed, AID-A2 #1-GFP, which is cytoplasmic in the presence of leptomycin B, was able to diffuse into the nucleus when the last 17 amino acids of AID were deleted (Fig. 6 b). The differential effect between treating AID-A2 #1-GFP with leptomycin B (see Fig. 5 c) and deleting the AID C-terminal domain suggests that this region is involved in cytoplasmic retention, apart from its role in nuclear export. This is reinforced by the observation that several nested N-terminal truncations of AID were able to confer preferentially cytoplasmic localization to GFP in leptomycin B treated 293T and Hela cells, including a fragment only encompassing residues 164-198 of AID (Fig. 6 c and Supplementary Fig. 7). The retention was not as strong as for full-length constructs, suggesting additional, perhaps structural, requirements.

We further used AID-A2 #1-GFP in an effort to find point mutations that could distinguish cytoplasmic retention from nuclear export. Affecting export but not retention would be expected to result in nuclear

exclusion of AID-A2 #1-GFP both in steady state and after leptomycin B. Mutating the NES residue L198S, produced such a phenotype (Fig. 6 d). The exclusively nuclear localization and loss of CRM1 binding of AID L198S-GFP (Fig. 6 d an e) indicated that L198S was largely causing export deficiency. Thus, the still preferentially cytoplasmic localization of AID-A2 #1 L198S-GFP suggests that export and retention can be functionally separated. Conversely, affecting retention but not export should result in nuclear exclusion of AID-A2 #1-GFP in steady state but allow diffusion after leptomycin B. We tested several mutations within the C-terminal domain of AID including at Tyr184, which could modulate retention through phosphorylation[20,21]. Most variants remained cytoplasmic in the context of AID-A2 #1-GFP and shuttled normally in the context of AID-GFP, except for those affecting residues Asp187 and Asp188 (Supplementary Table 1). In Hela cells, AID-A2 #1 D188A-GFP was able to diffuse into the nucleus only after leptomycin B treatment and AID-GFP bearing D187A or D188A were no longer excluded from the nucleus in steady state (Fig. 6 d). AID D188A-GFP was still largely cytoplasmic in 293T cells, in line with the differences observed between these two cell lines, but the double mutation D187A/D188A caused consistent loss of nuclear exclusion in HeLa and 293T cells (Fig. 6 d and Supplementary Fig. 7), suggesting that both contribute to retention. In addition, two further evidences support that mutations in Asp187 and As188 are preferentially affecting retention rather than export. First, AID D188A and AID D187A/D188A coimmunoprecipiated with CRM1 as efficiently as AID, in stark contrast with AID L198S, which did not (Fig. 6 e). Second, AID D188A-GFP accumulated faster than AID-GFP in the nucleus of leptomycin B-treated 293T cells, as it would be expected for a variant with reduced retention (Fig. 6 f and Supplementary Fig. 7).

Functional impact of altering AID localization balance

The efficiency of nuclear import and cytoplasmic retention would be predicted to influence AID biology. We performed CSR assays on *Aicda-/-* mouse B cells complemented by retroviral delivery with AID variants showing compromised nuclear import. Flag-AID was able to rescue isotype switching to IgG1 but less efficiently than untagged AID (Fig. 4 a). This reduction is more pronounced when the >2-fold higher protein expression of Flag-AID than AID is factored in (Fig. 7 b). On the other hand, while AID-GFP restored isotype switching, GFP-AID was completely unable to do so despite similar expression levels (Fig. 7 a and b). Thus, the effect of N-terminal tags on import seems proportional to their size, in keeping with the proposed masking of the NLS and implying that decreased nuclear import can limit CSR. We tested other AID variants with mutated import and cytoplasmic retention determinants by this assay. AID FYRN and AID-A2 50–54 were catalytically inactive (Fig. 7 c) and could not rescue CSR. AID with single point mutations at Arg50 or Asn51 have

been found to be CSR-proficient[39] but our constructs bear additional mutations, which most likely explains the difference. On the other hand, AID-A2 19–22 and 34–36 as well as AID D187A/D188A could not complement CSR despite being catalytically active (Fig. 7 a and c). The phenotype of AID-A2 19–22 and 34–36 may suggest that nuclear import is necessary for CSR, just as nuclear export seems to be[40]. However, it cannot be excluded that these chimeras have lost binding affinity to relevant AID partners[35]. Mutations at Asp187 and Asp188 may affect the C-terminal domain of AID that is required for CSR[41,42].

AID is rapidly degraded in the nucleus[19], so altering the balance between the mechanisms determining its subcellular localization should influence its half-life. If cytoplasmic retention were important in determining the subcellular localization of AID in steady state, reduced retention of AID should lead to an increased proportion of nuclear protein, resulting in shorter protein half-life. The degradation kinetics of AID D187A/D188A-GFP in DT40 B cells after cycloheximide treatment show that this is the case (Fig. 7 d). Since AID D187A/D188A-GFP has similar catalytic activity to AID-GFP, structural problems are unlikely to cause the destabilization. Importantly, after leptomycin B treatment the degradation of AID D187A/D188A-GFP was still more rapid than for AID-GFP (Fig. 7 d), as faster nuclear import would predict.

Finally, to examine the importance of cytoplasmic retention we determined the subcellular localization of endogenous AID in Ramos B cells in steady state and after leptomycin B by immunofluorescence. We could not detect any major increase in nuclear AID signal after incubation with leptomycin B (Fig. 7 e). This observation further indicates that nuclear export is not the only mechanism excluding AID from the nucleus and suggests the relevance of cytoplasmic retention for the compartmentalization of AID.

DISCUSSION

The subcellular localization of AID determines how much AID becomes in contact with the genome, probably a crucial parameter in balancing antibody diversification and off-target mutations/translocations. Nuclear export is one mechanism regulating AID localization[17,18,24]. We now provide strong evidence of two additional mechanisms influencing human AID subcellular distribution: active nuclear import and cytoplasmic retention.

Active nuclear import of AID

We demonstrate herein that AID is actively imported into the nucleus. Passive diffusion of AID into the nucleus was previously proposed based on consideration on AID size and the inability of AID fused to the C-terminus of a large protein to enter the nucleus[18,24]. The observation that fusions to the N-terminus of AID largely block its nuclear import, can now explain this difference. It has also been proposed that AID residues 1-26 contain a classical bipartite NLS[17]. However, the minimal AID fragment mediating nuclear accumulation encompasses from the N-terminus to somewhere between residues 160 and 181, pointing to the importance of protein conformation. The N-terminal domain contributes at least three of the four non-contiguous, positively charged determinants that we propose conform the AID NLS. Binding to importin-α adaptors, which mediate the nuclear import of proteins with classical basic NLS[29,30,36], is in line with this. Without the 3D structure of an AID-importin complex it is difficult to assign which residues make direct contact and which play a structural role in displaying those residues. Hydrophobic residues W20, V18 and Y13, previously identified as necessary for AID nuclear accumulation[43], may be examples of the latter kind. NLSs often overlap with nucleic acid binding domains[44]. Indeed, Arg24 and Arg112 might be involved in DNA binding by comparison with APOBEC3G[31]. However, three import-deficient but catalytically active AID-A2 chimeras (#5, 19–22 and 34–36) provide separation of function between active import and nucleic acid binding. Similarly to AID, the N-terminal domain of APOBEC1 is necessary but not sufficient for nuclear import[45,46] suggesting that APOBEC1 might also have a conformational NLS.

The stoichiometry and architecture of the biologically relevant AID molecule is unknown but many indirect evidences suggest that it will have quaternary structure[33,35,42,47,48] as all cytidine deaminases including the APOBECs do[33,49,50]. Predicting AID dimerization through β2 (ref. 33) produces a thermodynamically robust model, which is supported by the compromised oligomerization of AID FYRN. Since AID FYRN is also compromised for nuclear import it suggests a role for AID dimerization in conforming the NLS, which is also consistent with the location of the import

determinants in our AID dimer model (Fig. 4 a). However, AID FYRN is unstable and inactive. Although this would not be unexpected if a dimer was the minimal functional unit of AID[33], we cannot rule out other structural defects impinging on the localization of AID FYRN. Solving this would need elucidating the structure of AID by more sophisticated techniques than coimmunoprecipitation, exceeding the focus of this work. In any case, AID FYRN still supports the conformational NLS and the importance of the tertiary structure for nuclear import.

Cytoplasmic retention of AID

We show in a number of ways that AID is unable to diffuse into the nucleus when both active import and export are inhibited, strongly suggesting the existence of cytoplasmic retention. The C-terminal domain of AID, which we propose to be a flexible helix, seems necessary for cytoplasmic retention. This domain also bears the NES[17,18,24], complicating the mapping of cytoplasmic retention determinants, which may partially overlap with it. However, comparison of mutations L198S and D187A/D188A provides good evidence that export and retention are distinct and can be separated. Structural integrity of AID may also be important as suggested by the less strong retention of some N-terminal truncations of AID and by the apparently reduced retention of AID-FYRN, which unlike AID is not able to outcompete a heterologous NLS (Supplementary Fig. 5). Coincidently, a structural determinant mediates cytoplasmic retention of APOBEC3G[49,51].

Regarding the identity of a putative AID cytoplasmic anchor, the observation that 3XFlag-AID was resistant to digitonin washing from the cytoplasm, led others to propose retention of AID mediated by binding to tubulin[52]. We did not detect any change in AID-GFP localization after treatment with nocodazole (not shown) but other cytoskeleton molecules might be involved and further research is needed to settle this issue. Other candidates to mediate AID cytoplasmic retention may be the 14-3-3 proteins given that AID is phosphorylated[29]. The analysis of partners in the uncharacterized AID high molecular weight cytoplasmic complexes (K. Xue and M. Neuberger personal communication); or comparisons between HeLa and 293T cells, which show different kinetics for AID nuclear import, may help to identify relevant cytoplasmic retention and/or nuclear import AID partners.

Functional implications

Our findings demonstrate that AID compartmentalization is not determined just by nuclear export, as the available evidences indicated so far[18,24]. We show that AID passive diffusion from the cytoplasm into the nucleus is hindered and therefore AID needs to be actively imported. Both, cytoplasmic retention and nuclear import influence the biology of AID. Hindering import results in cytoplasmic

retention and decreased CSR. AID variants with reduced cytoplasmic retention, are no longer excluded from the nucleus in steady state resulting in shorter protein half-life. Thus, nuclear export seems insufficient to outweigh AID nuclear import. The observation that leptomycin B treatment of Ramos cells results in little AID nuclear translocation supports this scenario, although this should be confirmed in non-transformed B cells. The difference with AID-GFP, which accumulates in the nucleus of Ramos, could be explained by a combination of overexpression and the C-terminal tag weakening retention, just as an N-terminal tag affects import. AID is vastly confined to the cytoplasm of germinal center B cells[48,53-55]. Only a very small proportion of the B cells in rabbit appendix[48] and human tonsils[53], and hardly any cell in unsynchronized populations overexpressing AID-GFP, seem to contain predominantly nuclear AID. In the light of our results it seems unlikely that such accumulation might be cell cycle regulated or has much physiological significance. We would rather suggest a model in which a small proportion of nuclear AID exists at any given time, resulting from the competition between active import, nuclear export and cytoplasmic retention. This steady state level of nuclear AID might be enough for efficient antibody gene diversification as suggested by the CSR levels elicited by apparently exclusively cytoplasmic AID variants, like AID V18S/R19V[43] or Flag-AID. Perturbing any of the mechanisms determining AID subcellular localization will affect this equilibrium and impact on the efficiency and specificity of antibody diversification.

Acknowledgments

Author contributions

AMP, AO, YH, VAC, AB and JMD performed research. All authors analyzed data and discussed results. JMD designed research and wrote the paper.

Figure Legends

Figure 1 AID nuclear import is an active process. (**a**) Representative confocal images showing the kinetics of AID-GFP nuclear accumulation in stably transfected Ramos and transiently transfected HeLa and 293T cells. Time in h after addition of leptomycin B is indicated on the left. Propidium iodide (PI) staining of nuclei is included for Ramos cells. (**b**) Energy depletion experiments in HeLa cells transiently expressing NLS^{SV40}-GFP or the constitutively nuclear AID L198S-GFP. Representative confocal images are shown for untreated cells (Ctrl), energy-depleted cells (ΔGlc), or energy-depleted cells reincubated in complete medium (ΔGlc+Glc). The average proportion of cells + *s.e.m.* with clear cytoplasmic localization is plotted for each condition as determined from 2 independent experiments (58-100 cells scored). Student t-test *p* values are shown. (**d**, **e**) – Representative confocal images of 293T cells transiently expressing the indicated fusions proteins. PI was used to stain the nuclei. β-Lac, β-Lactamase; β-Gal, β-Galactosidase; UNG2, nuclear isoform of mouse uracil DNA-glycosylase. The punctuated nature of the signal for the β-Gal fusion proteins in the nucleus was consistently observed in all cells for both constructs. All images are at 630X magnification. Bar, 10 μm.

Figure 2 Most of the AID protein is required to mediate efficient nuclear import (**a**) Schematic representation of C-terminally GFP-tagged AID truncations (named after the last AID residue included) and their subcellular localization in transiently transfected 293T cells. Cells were scored from confocal images (representative pictures shown in supplementary figure 3) and classified according to the predominant (*i.e.* observed in >85% of the cells) subcellular localization (Loc) as cytoplasmic (C), throughout the cell (N+C) or nuclear (N). (**b**) Representative confocal images of HeLa cells in steady state transiently expressing the indicated constructs described in **a**. The localization of GFP control and full length AID after leptomycin B treatment are shown for comparison. All transfections were performed in parallel and images acquired using the same settings. (**c**) Representative confocal images of Ramos cells in steady state transiently expressing the indicated constructs described in **a**. Nuclear staining with propidium iodide (PI) is shown. (**d**) Representative confocal images of 293T cells in steady state transiently expressing the indicated constructs with a C-terminal GFP tag. NLS^{NP}, bipartite nuclear localization signal from nucleoplasmin; β-Gal, β-Galactosidase. (**e**) Schematic representation of AID-A2 chimeric proteins. Full length AID and A2 are drawn to scale and aligned. In each AID-A2 chimera (#1 to #5) the indicated AID amino acid range was replaced with the homologous region from

15

A2. NES, nuclear export signal; K/R-rich, AID N-terminal domain rich in Lys and Arg. Representative images of HeLa cells transiently expressing the indicated C-terminally GFP tagged proteins in steady state and after leptomycin B treatment (+ Lept B) are shown. All images are at 630X magnification. Bars, 10 μm.

Figure 3 Several non-contiguous, positively charged residues are critical for AID nuclear import. (**a**) AID oligomerization assay. The physical association between C-terminally Flag- and GFP-tagged versions of the indicated proteins was monitored by detecting the GFP-tagged proteins in western blots of anti-Flag immunoprecipitates from transiently cotransfected 293T cell extracts. The filters were subsequently probed with anti-Flag to confirm similar immunoprecipitation of the bait. Aliquots representing 5 % of the extracts were probed to control for expression levels of the GFP-tagged proteins. One representative of three experiments done is shown. (**b**) Representative confocal images of 293T cells transiently expressing the indicated fusion proteins. β-Gal, β-galactosidase. Leptomycin B was added where indicated. (**c**) Replacements of 3-5 consecutive residues in the N-terminal region of AID with the corresponding residues in A2. Identities between AID and APOBEC2 are indicated in bold on the AID sequence. Representative confocal pictures of HeLa cells expressing each AID-A2-GFP protein are shown in the steady state or after leptomycin B treatment. (**d**) Representative confocal images of HeLa cells transiently expressing AID-A2 chimeras in which the two indicated residue stretches of AID were simultaneously replaced. (**e**) Representative confocal pictures of HeLa cells transiently expressing untagged AID and AID-A2 #1 and 34-36. AID was detected by immunofluorescence with anti-AID followed by biotinylated anti-rat and anti-biotin FITC in steady state or after 4 h of leptomycin B treatment.. In **c-d** the proportion of cells with largely cytoplasmic (black), homogeneous throughout the cell (grey) or nuclear (white) GFP signal was scored from >6 random fields from two independent experiments and the proportions indicated in the stacked bars plots below each corresponding picture. Images at 630X except panel d at 400X magnification. Bars, 10 μm.

Figure 4 AID conformational NLS and interaction with importin-α. (**a**) Detailed location in a putative AID dimer (monomers distinguished in green and pink) of those residues whose mutation, either alone or as part of an amino acid range in AID-A2 chimeras, was found to affect the efficiency of AID nuclear import (blue sticks). Dimerization through β2 as in APOBEC2[33] was assumed based on similarity and available biochemical data (see supplementary methods). For clarity, not all mutated

16

residues are labeled. Asterisks indicate residues belonging to the second monomer. Tyr13, Val18 and Trp20, identified elsewhere[43], are also included. (**b**) Predicted solvent accessible surface of the putative AID dimmer with those residues identified in **a** shown in shades of blue. Positively charged residues within the import determinants are highlighted in bright blue to show the bias for exposed basic amino acids. (**c**) Glutathione-sepharose beads loaded with the indicated purified GST-fusion protein were used in pull down assays on cell extracts from Ramos B cells stably expressing AID-Flag. The material eluted from the beads was analyzed by western blot using anti-AID. Ponceau-S staining of the transferred proteins is shown as an indication of GST-protein precipitation. (**d**) Pull down of the indicated GFP-tagged proteins from transiently transfected 293T cell extracts using purified GST of GST-importin-α3 immobilized on glutathione-agarose. The material eluted from the beads and a 2% aliquot of the input lysates were analyzed by western blot using anti-GFP. Ponceau staining of the transferred proteins is shown as an estimation of similar GST-protein precipitation. Note that AIDΔN26 and AID 181 were appreciably degraded in the extracts (AID 181 was not visible in the extracts but it was after concentration by the pull down).

Figure 5 AID cannot passively diffuse into the nucleus. (**a**) Representative confocal images of Ramos B cells transiently expressing AID-A2 #1-GFP or APOBEC2-GFP after leptomycin B treatment. Both proteins show the same distribution in steady state (not shown). (**b**) Representative confocal images of Ramos B cells stably expressing GFP-AID (top) and DT40 B cells stably expressing Flag-AID (bottom) in steady state of after leptomycin B treatment. Flag was detected by immunofluorescence using mouse anti-Flag followed by anti-mouse AlexaFluor488. Nuclei are shown by PI staining. (**c**) Stacked bars plot comparing the subcellular localization of AID-GFP, GFP-AID and AID-A2 #1-GFP in Hela cells at different time points after addition of leptomycin B. The number (*n*) of cells indicated above each bar was scored from multiple fields and the proportion of cells at each time point showing largely cytoplasmic (black), homogeneous (grey) or nuclear (white) GFP signal determined. Representative pictures are shown in supplementary figure 6. (**d**) Representative confocal pictures of Hela cells transiently expressing AID-GFP or A2-GFP pretreated or not for 1 h with H_2O_2 before treating for 2 h with leptomycin B where indicated. Subcellular localization was scored and plotted as in **c** from three independent experiments. (**e**) Representative pictures of a similar experiment to that in **d** but using untagged AID, which was detected by immunofluorescence as in figure 3. Subcellular localization was scored and plotted as in **c** from two independent experiments. All images are at 630X except **d** and **e** at 400X magnification. Bars, 10 μm.

Figure 6 The C-terminal domain of AID contains a cytoplasmic retention determinant. (**a**) Representative confocal images of HeLa cells transiently expressing the indicated AID-A2 chimeras (schematized in Fig. 2) with N-terminal GFP tag after leptomycin B treatment. (**b**) Representative confocal images of HeLa cells transiently expressing AID-A2 #1-GFP in the presence of leptomycin B or carrying a deletion of the last 17 residues of AID. (**c**) Representative confocal images of HeLa cells treated with leptomycin B, transiently expressing GFP, A2 70-224-GFP (APOBEC2 with a deletion of the first 69 residues) and AID 55-198-GFP (AID with a deletion of the first 54 residues). These truncations were designed so that the resulting fragments are structurally equivalent. Nuclei are shown by propidium iodide staining (PI). Experiments including a whole set of AID N-terminal truncations in Hela and 293T cells is shown in supplementary figure 7. (**d**) Representative confocal images of HeLa cells transiently expressing C-terminally GFP-tagged constructs of AID or AID-A2 #1 with the indicated point mutations. Cells were treated with leptomycin B where indicated. In panels **a**, **b**, **c** and **d** the subcellular localization of the constructs within the cells was scored as predominantly cytoplasmic (black), homogenous throughout the cell (grey) or largely nuclear (white) from the indicated number of cells (*n*) and the proportions plotted in the stacked bars below each corresponding picture. All images are at 630X magnification. Bars, 10 μm. (**e**) 293T cells overexpressing untagged full-length human CRM1 were cotransfected with the indicated AID-Flag variants. AID was immunoprecipitated by anti-flag beads and the presence of CRM1 in the precipitate analyzed by western blot. The IP filters were subsequently probed with anti-flag HRP to confirm precipitation of the bait. One of two experiments yielding identical results is shown. (**f**) Line plot depicting the import kinetics of AID-GFP vs AID D188A-GFP in transiently transfected 293T cells. The proportion of cells with exclusively nuclear GFP signal was scored at each time point after the addition of leptomycin B to the cultures. Two independent experiments are plotted, which only had partially overlapping time points. For $t = 0$, 1 and 2 h the average and *s.e.m* of both experiments is plotted, the other points represent one experiment only. Representative pictures are shown in supplementary figure 7.

Figure 7 Altering the balance between AID subcellular localization mechanisms has functional consequences. (**a**) B cells from AID-deficient mice were infected with retroviruses encoding untagged AID or Flag-AID in pMX-ires-GFP or GFP control, GFP-AID or the indicated AID-GFP variants in pMX. The proportion of infected (GFP positive) cells that have switched to IgG1 is indicated in the top right quadrant of representative flow cytometry profiles and summarized in the bar graphs. Different

symbols indicate cells coming from individual mice. Filled and empty symbols indicated independent experiments. (**b**) Western blots comparing the expression of the indicated proteins. PCNA was used as loading control. (**c**) *E coli* mutation assay to monitor the catalytic activity of AID-GFP and the indicated derived mutants. Bottom plot is in log scale. Combined data from two independent experiments are plotted for each construct. (**d**) DT40 B cells stably expressing GFP control, AID-GFP or AID D187A/D88A-GFP (AID DD-GFP) were pretreated with cycloheximide for 1 h and the GFP signal monitored over time by flow cytometry in the presence or absence of leptomycin B. For each construct and condition, the mean fluorescent intensity (MFI) of the GFP signal at each time point was normalized to the MFI at $t = 0$. The averages $\pm$ *s.d.* of three experiments are plotted. The two curves done in the presence of leptomycin B were analyzed by t-test and found to be significant with *p<0.05* at $t = 1$ and *p<0.001* at all other time points. (**e**) Representative confocal images of untransfected Ramos B cells treated or not with leptomycin B for 2 or 5 h (and where indicated with 10 μM MG132 proteasome inhibitor) stained with anti-AID followed by biotinylated anti-Rat and anti-biotin FITC. Staining in which primary (1ry) anti-AID was omitted or using a Ramos clone with low AID expression (AID$^{\text{lo}}$)[56] were used as specificity controls. Western blot analysis comparing the expression of AID in Ramos and Ramos AID$^{\text{lo}}$ are shown on the right.

METHODS

DNA constructs. N-terminally GFP-tagged human AID, APOBEC2 and the AID-APOBEC2 chimeras in pEGFP-C3 were a gift of Dr S. Conticello[35]. We generated C-terminally GFP-tagged versions by PCR amplification of the open reading frames using oligonucleotides OJ60 and OJ166 for AID, OJ196 and OJ197 for APOBEC2 and OJ60-OJ166 for all AID-A2 except #5, for which OJ60-OJ197 were used (see supplementary methods for all for oligonucleotide sequences). The 5' oligonucleotides introduced a BamHI site and a CAA triplet, a "cold" translation initiation context to reduce expression levels, just before the ATG. The 3' oligonucleotides eliminated the stop codon and introduced an EcoRI site at the end of the coding regions. All fragments were cloned into BglII/EcoRI-digested pEGFP-N3. The AID protein truncations used were generated by PCR amplification using the indicated oligonucleotides (supplementary methods) and identically cloned into pEGFP-N3. We introduced point mutations by the quickchange method using the oligonucleotides listed in supplementary methods. We cloned the open reading frames of β-Lactamase, amplified from pUC18 using OJ213 and OJ214, and β-Galactosidase, amplified from pIND/lacZ (Invitrogen) using OJ172 and OJ174, as EcoRI-BamHI fragments into pEGFP-N3. To construct C-terminally flag-tagged versions of some of the proteins, EGFP was excised from pEGFP-N3 using EcoRI and NotI and replaced by a synthetic Flag/HA-tag (oligonucleotides OJ215 and OJ216). The nuclear localization signals from the SV40 large T antigen and from nucleoplasmin were constructed by annealing oligonucleotides OJ190/OJ191 and OJ192/OJ193, respectively. We used pCDNA3.1 (invitrogen) to express untagged AID and AID mutants. Construct names throughout the manuscript indicate the actual order of the fragments in the fusion proteins.

Cell methods and microscopy. Transfections were done as indicated in supplementary methods. We inhibited nuclear export by using leptomycin B (LC labs, Woburn, MA) at 50 ng ml^{-1} for 4 h for 293T cells, 50 ng ml^{-1} for 2 h for Hela cells and 10 ng ml^{-1} for 2 h for Ramos and DT40 B cells, unless indicated differently. We inhibited nuclear import by pretreating the cells for 1 h with freshly diluted 10 mM H_2O_2 before any further treatment or analysis. Adherent cells were grown on coverslips and B cell lines were allowed to attach to poly-L-Lysine-coated coverslips for 20 min at 37°C in PBS. For immunofluorescence the fixed cells were permeabilized in 0.25% (v/v) Triton X-100, blocked overnight and stained in PBS 5% (w/v) BSA using mouse MAb anti-Flag M2 (1:100 Sigma-Aldrich) followed by anti-mouse AlexaFluor 488 of 546 (1:1000 Invitrogen) or rat MAb anti-AID (1:100 EK2 5G9, Cell Signaling) followed by biotinylated goat anti-Rat antibodies (1:100 Abcam) detected with anti-biotin FITC (1:50 Milteny Biotec). For energy depletion experiments we plated transfected cells in

triplicate. We subsequently incubated two of the wells in glucose-free DMEM (Wisent Inc) supplemented with 10% (v/v) dialyzed FCS, 6 mM 2-deoxy-D-glucose (Sigma-Aldrich), to inhibit glycolysis, and 10 mM sodium azide (Bioshop), to inhibit mitochondrial respiration, for 3 h. After microscopic examination we fixed one of the wells and incubated the remaining well for further 3 h in complete DMEM before fixing and processing along with the untreated control. For microscopy we fixed cells in 3.7% (w/v) p-formaldehyde for 15 min, washed 3 times in PBS and stained 15 min in PBS 0.5% (v/v) Triton-X100, 200 μg ml^{-1} RNase A, 10 μg ml^{-1} propidium iodide; washed in PBS and mounted using FluorSafe (Calbiochem). Detailed confocal image acquisition and scoring is explained in supplementary methods.

Coimmunprecipitation, 293T cells cotransfected with GFP and Flag-tagged versions of the indicated proteins were homogenized in Lysis Buffer [20 mM Tris pH 7.8, 137 mM NaCl, 10 % (v/v) Glycerol, 2mM EDTA, 1 % (v/v) Triton X-100, Complete (Roche) protease inhibitors] 48 h post-transfection. Lysates were incubated with 20 μl of anti-Flag®M2 affinity gel (Sigma-Aldrich) for 3 h at 4°C and eluted with 50 μL of 3x Flag peptide (Sigma-Aldrich) following the manufacturer protocol. Eluates were analyzed by western blot with anti-eGFP and 1:5000 anti-Flag-HRP (Sigma-Aldrich) or 1:1000 anti-CRM1 (Santa Cruz Biotechnology). Full-length human CRM1 cloned into pCDNA3 was a gift from Dr S Swaminathan. GST-fusion protein pull downs were done on cell extracts using loaded glutathione-agarose beads as detailed in supplementary methods. Full-length human importin-α3 cloned in pGEX4T1 was a gift of Dr I. Macara, a similar construct with human importin-α5 was a gift of Dr M. Malim. Full-length human importin-α1 was subcloned into pGEX-4T1 as an EcoRI/XhoI fragment from pJG4-5-importin-α1[45], a gift of Dr N. Navaratnam. Western blots were developed using SuperSignal® West Pico Chemiluminiscent substrate (Thermo Scientific).

AID assays. We assayed catalytic activity of AID by a mutation assay as described[57] using the ung-deficient *E coli* strain BW310 expressing AID-GFP and various mutants subcloned as NheI-NotI fragments into a modified pTrc99a. We assayed isotype class switching by retrovirally deliverying AID variants into *in vitro* activated B cells, obtained from *Aicda-/-* mice. The IRCM Animal Ethics Committee approved all animal handling. Purification of resting B cells from splenic lymphocytes by CD43-depletion, retroviral infection and analysis of isotype class switching to IgG1 has been described[58]. AID-GFP variants and GFP-AID were subcloned as EcoRI-NotI fragments into the pMXs retroviral vector. Untagged human AID and Flag-AID were subcloned as BamHI-NotI fragments into pMXs-ires-GFP[59]. AID stability monitoring is explained in supplementary methods.

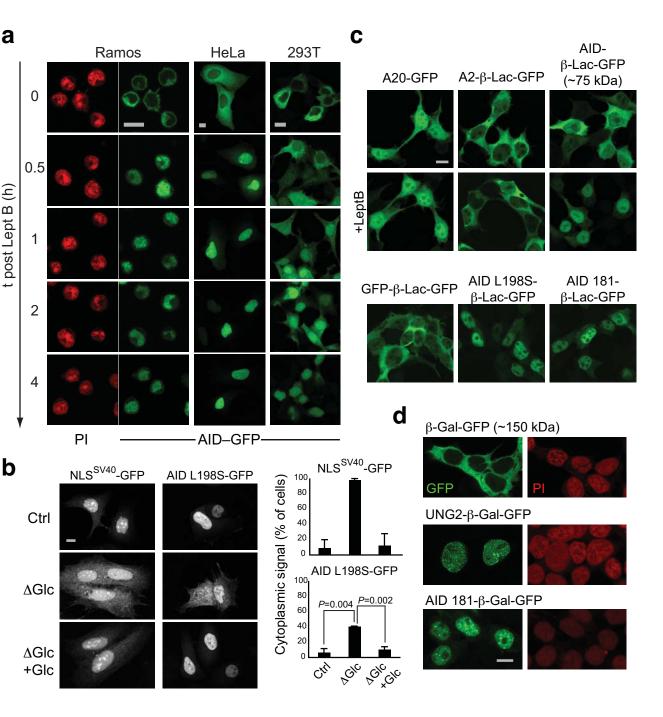**Structure modeling**. See supplementary methods.

Satistics. Where indicated, two experimental data sets were compared using paired, one-tailed, Student's t-Test.

References

1.   Muramatsu, M. et al. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553-63 (2000).
2.   Revy, P. et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* **102**, 565-75 (2000).
3.   Peled, J.U. et al. The Biochemistry of Somatic Hypermutation. *Annu Rev Immunol* (2007).
4.   Di Noia, J.M. & Neuberger, M.S. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem* **76**, 1-22 (2007).
5.   Chaudhuri, J. et al. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv Immunol* **94**, 157-214 (2007).
6.   Martin, A. & Scharff, M.D. Somatic hypermutation of the AID transgene in B and non-B cells. *Proc Natl Acad Sci USA* **99**, 12304-8 (2002).
7.   Okazaki, I.M. et al. Constitutive expression of AID leads to tumorigenesis. *J Exp Med* **197**, 1173-81 (2003).
8.   Liu, M. et al. Two levels of protection for the B cell genome during somatic hypermutation. *Nature* **451**, 841-5 (2008).
9.   Pasqualucci, L. et al. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc Natl Acad Sci U S A* **95**, 11816-21 (1998).
10.   Shen, H.M., Peters, A., Baron, B., Zhu, X. & Storb, U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science* **280**, 1750-2 (1998).
11.   Dorsett, Y. et al. A role for AID in chromosome translocations between c-myc and the IgH variable region. *J Exp Med* **204**, 2225-32 (2007).
12.   Ramiro, A.R. et al. AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* **118**, 431-8 (2004).
13.   Crouch, E.E. et al. Regulation of AID expression in the immune response. *J Exp Med* **204**, 1145-56 (2007).
14.   de Yébenes, V.G. et al. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J Exp Med* (2008).
15.   Dorsett, Y. et al. MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity* **28**, 630-8 (2008).
16.   Teng, G. et al. MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity* **28**, 621-9 (2008).
17.   Ito, S. et al. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc Natl Acad Sci USA* **101**, 1975-80 (2004).
18.   McBride, K.M., Barreto, V.M., Ramiro, A.R., Stavropoulos, P. & Nussenzweig, M.C. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J Exp Med* **199**, 1235-44 (2004).
19.   Aoufouchi, S. et al. Proteasomal degradation restricts the nuclear lifespan of AID. *J Exp Med* **205**, 1357-68 (2008).
20.   Basu, U. et al. The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature* **438**, 508-11 (2005).
21.   McBride, K.M. et al. Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. *Proc Natl Acad Sci U S A* **103**, 8798-803 (2006).
22.   Pasqualucci, L., Kitaura, Y., Gu, H. & Dalla-Favera, R. PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. *Proc Natl Acad Sci USA* **103**, 395-400 (2006).

23.   Rada, C., Jarvis, J.M. & Milstein, C. AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc Natl Acad Sci USA* **99**, 7003-8 (2002).

24.   Brar, S.S., Watson, M. & Diaz, M. Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. *J Biol Chem* **279**, 26395-401 (2004).

25.   Richardson, W.D., Mills, A.D., Dilworth, S.M., Laskey, R.A. & Dingwall, C. Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores. *Cell* **52**, 655-64 (1988).

26.   Newmeyer, D.D. & Forbes, D.J. Nuclear import can be separated into distinct steps in vitro: nuclear pore binding and translocation. *Cell* **52**, 641-53 (1988).

27.   Guiochon-Mantel, A. et al. Nucleocytoplasmic shuttling of the progesterone receptor. *EMBO J* **10**, 3851-9 (1991).

28.   Larijani, M. et al. AID associates with single-stranded DNA with high affinity and a long complex half-life in a sequence-independent manner. *Mol Cell Biol* **27**, 20-30 (2007).

29.   Macara, I.G. Transport into and out of the nucleus. *Microbiol Mol Biol Rev* **65**, 570-94, table of contents (2001).

30.   Görlich, D. & Kutay, U. Transport between the cell nucleus and the cytoplasm. *Annu Rev Cell Dev Biol* **15**, 607-60 (1999).

31.   Chen, K. et al. Structure of the DNA deaminase domain of the HIV-1 restriction factor APOBEC3G. *Nature* **452**, 116-9 (2008).

32.   Holden, L.G. et al. Crystal structure of the anti-viral APOBEC3G catalytic domain and functional implications. *Nature* **456**, 121-4 (2008).

33.   Prochnow, C., Bransteitter, R., Klein, M.G., Goodman, M.F. & al., e. The APOBEC-2 crystal structure and functional implications for the deaminase AID. *Nature* **445**, 447-51 (2006).

34.   Nilsen, H. et al. Analysis of uracil-DNA glycosylases from the murine Ung gene reveals differential expression in tissues and in embryonic development and a subcellular sorting pattern that differs from the human homologues. *Nucleic Acids Res* **28**, 2277-85 (2000).

35.   Conticello, S.G. et al. Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNBL1. *Mol Cell* **31**, 474-84 (2008).

36.   Lange, A. et al. Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem* **282**, 5101-5 (2007).

37.   Miyamoto, Y. et al. Cellular stresses induce the nuclear accumulation of importin alpha and cause a conventional nuclear import block. *J Cell Biol* **165**, 617-23 (2004).

38.   Kodiha, M., Chu, A., Matusiewicz, N. & Stochaj, U. Multiple mechanisms promote the inhibition of classical nuclear import upon exposure to severe oxidative stress. *Cell Death Differ* **11**, 862-74 (2004).

39.   Shivarov, V., Shinkura, R. & Honjo, T. Dissociation of in vitro DNA deamination activity and physiological functions of AID mutants. *Proc Natl Acad Sci USA* **105**, 15866-71 (2008).

40.   Doi, T. et al. 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 USA* (2009).

41.   Barreto, V., Reina-San-Martin, B., Ramiro, A.R., McBride, K.M. & Nussenzweig, M.C. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol Cell* **12**, 501-8 (2003).

42.   Ta, V. et al. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat Immunol* **4**, 843-8 (2003).

43.   Shinkura, R. et al. Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat Immunol* **5**, 707-12 (2004).

44.   LaCasse, E.C. & Lefebvre, Y.A. Nuclear localization signals overlap DNA- or RNA-binding domains in nucleic acid-binding proteins. *Nucleic Acids Res* **23**, 1647-56 (1995).

45.   Chester, A. et al. The apolipoprotein B mRNA editing complex performs a multifunctional cycle and suppresses nonsense-mediated decay. *EMBO J* **22**, 3971-82 (2003).

46.   Yang, Y. & Smith, H.C. Multiple protein domains determine the cell type-specific nuclear distribution of the catalytic subunit required for apolipoprotein B mRNA editing. *Proc Natl Acad Sci USA* **94**, 13075-80 (1997).

47.   Dickerson, S.K., Market, E., Besmer, E. & Papavasiliou, F.N. AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med* **197**, 1291-6 (2003).

48.   Yang, G. et al. Activation-induced deaminase cloning, localization, and protein extraction from young VH-mutant rabbit appendix. *Proc Natl Acad Sci USA* **102**, 17083-8 (2005).

49.   Bennett, R.P., Presnyak, V., Wedekind, J.E. & Smith, H.C. Nuclear Exclusion of the HIV-1 Host Defense Factor APOBEC3G Requires a Novel Cytoplasmic Retention Signal and Is Not Dependent on RNA Binding. *J Biol Chem* **283**, 7320-7 (2008).

50.   Lau, P.P., Zhu, H.J., Baldini, A., Charnsangavej, C. & Chan, L. Dimeric structure of a human apolipoprotein B mRNA editing protein and cloning and chromosomal localization of its gene. *Proc Natl Acad Sci USA* **91**, 8522-6 (1994).

51.   Stenglein, M.D., Matsuo, H. & Harris, R.S. Two regions within the amino-terminal half of APOBEC3G cooperate to determine cytoplasmic localization. *J Virol* **82**, 9591-9 (2008).

52.   Wu, X., Geraldes, P., Platt, J.L. & Cascalho, M. The double-edged sword of activation-induced cytidine deaminase. in *J Immunol* Vol. 174 934-41 (2005).

53.   Cattoretti, G. et al. Nuclear and cytoplasmic AID in extrafollicular and germinal center B cells. *Blood* **107**, 3967-75 (2006).

54.   Greiner, A. et al. Differential expression of activation-induced cytidine deaminase (AID) in nodular lymphocyte-predominant and classical Hodgkin lymphoma. *J Pathol* **205**, 541-7 (2005).

55.   Pasqualucci, L. et al. Expression of the AID protein in normal and neoplastic B cells. *Blood* **104**, 3318-25 (2004).

56.   Zhang, W. et al. Clonal instability of V region hypermutation in the Ramos Burkitt's lymphoma cell line. *Int Immunol* **13**, 1175-84 (2001).

57.   Petersen-Mahrt, S.K., Harris, R.S. & Neuberger, M.S. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**, 99-103 (2002).

58.   Di Noia, J.M. et al. Dependence of antibody gene diversification on uracil excision. *J Exp Med* **204**, 3209-19 (2007).

59.   Kitamura, T. et al. Retrovirus-mediated gene transfer and expression cloning: powerful tools in functional genomics. *Exp Hematol* **31**, 1007-14 (2003).

a

	Ramos	HeLa	293T
0			
0.5			
1			
2			
4			

t post Lept B (h)

PI ——————AID–GFP——————

b

NLS^{SV40}-GFP   AID L198S-GFP

Ctrl

ΔGlc

ΔGlc
+Glc

NLS^{SV40}-GFP

AID L198S-GFP

Cytoplasmic signal (% of cells)

P=0.004 P=0.002

Ctrl   ΔGlc   ΔGlc
+Glc

c

A20-GFP   A2-β-Lac-GFP   AID-
β-Lac-GFP
(~75 kDa)

+LeptB

GFP-β-Lac-GFP   AID L198S-
β-Lac-GFP   AID 181-
β-Lac-GFP

d

β-Gal-GFP (~150 kDa)

GFP   PI

UNG2-β-Gal-GFP

AID 181-β-Gal-GFP

Figure 1
(Di Noia)

Figure 2
(Di Noia)

a

WB

Anti-GFP

Anti-Flag

IP anti-Flag

AID-GFP
AIDF46A-GFP
AIDF46A/Y48A-GFP
AID FYRN-GFP
A2-GFP

—47

AID-Flag

—25

Anti-GFP

Lysates

—62
—47

GFP
AID-GFP
AID FYRN-GFP
A2-GFP

—47

AID FYRN-Flag

—25

—83
—62
—47

—32
—25

b

AID-GFP

+ Lept B

AID FYRN-GFP

+ Lept B

AID 181-βGal-GFP

Ctrl   FYRN

c

```
        19      25        34    39   43       50    55
AID ...KNVRWAKGRRETYLCYVVKRRDSATSFSLDFGYLRNKNG–CHVE...
A2        EYSS  NKTF     EAQ  GGQV QASR     EDEHAAAHA
```

AID-A2

| 19–22 | 25–28 | 34–36 | 39–42 | 43–46 | 50–54 | 55–57 |

+ Lept B

n = 50   n = 20   n = 44   n = 25   n = 19   n = 26   n = 19

■ C
▨ N+C
□ N

d

AID-A2
19–22/34–36

+ Lept B

n = 36

AID-A2
34–36/50–54

+ Lept B

n = 34

e

AID   + Lept B

n = 25   n = 22

AID-A2 #1

n = 16   n = 21

AID-A2 34–36

n = 12   n = 9

Figure 3
(Di Noia)

a

Arg19
Lys22
Lys52
Arg35*
Arg36*
Arg24
Arg112
Arg50
Arg34*
Tyr13
Trp20
β2 β2*

b

c

GST
GST-α1
GST-α3
GST-α5
Input

32
25

anti-Flag

87
62
47
32
25

Ponceau

GST
GST-α

d

AID
AID
A2
AID FYRN
AIDΔN
AID 181

62
47

anti-GFP

83
62
47
32
25

GST
GST-importin α3
Ponceau

62
47
32
25

Lysates

Figure 4
(Di Noia)

a

AID-A2 #1-GFP + Lept B   A2-GFP + Lept B

GFP   PI

b

GFP-AID   + LeptB (4 h)

GFP   PI

Flag-AID

α-FLAG   PI

c

AID-GFP   GFP-AID   #1-GFP

Cells (%)

■ C
▨ N+C
☐ N

n = 23 14 28 28 32   17 6 16 27 30   12 6 15 28 28

100
75
50
25
0

0 1 2 4 6

t post lept B (h)

d

AID-GFP   A2-GFP

n = 57 n = 53 n = 6

+ LeptB

n = 63 n = 49 n = 20

+ H_2O_2

e

Untagged AID   + H_2O_2

n = 17 n = 17

+ LeptB

n = 16 n = 18

Figure 5
(Di Noia)

a GFP-AID-A2 (+ Lept B)

#1 #2 #3 #4 #5

■C
▨N+C
□N

n = 50   *n* = 41   *n* = 16   *n* = 14   *n* = 28

b

AID-A2 #1-GFP   AID-A2 #1 ΔC17-GFP

+ LeptB

n = 74   *n* = 48

c

GFP   A2 70-224-GFP   AID 55-198-GFP

+ Lept B   PI   + Lept B   PI   + Lept B   PI

d

AID-A2 #1-GFP   AID-GFP

steady state   + Lept B   steady state   + Lept B

L198S
n = 15   *n* = 30   *n* = 26

D187A
n = 26   *n* = 20   *n* = 37   *n* = 24

D188A
n = 17   *n* = 28   *n* = 22   *n* = 16

D187A/D188A
n = 22   *n* = 19   *n* = 12   *n* = 15

e

AID-Flag/HA

vector wt D187A/D188A L196S/L198S D188A L198S WB

175 —   CRM1
25 —   Flag

IP anti-Flag

175 —
83 —   CRM1

Lysates

f

● D188A
■ wt

Cells with nuclear-only AID-GFP (%)

Time post lept B (min)

Figure 6
(Di Noia)

a

Ctrl 0% | uninfected 0% | GFP 0% | AID-GFP 14% | FYRN-GFP 1% | D188A-GFP 1%

AID 20% | Flag-AID 17% | GFP-AID 0% | AID DD-GFP 0% | 19-22-GFP 0% | 34-36-GFP 0% | 50-54-GFP 0%

IgG₁ ↑
GFP →

IgG1 (%)

P=0.007

Ctrl AID Flag-AID

pMX-iresGFP

GFP AID-GFP GFP-AID AID DD-GFP

pMX

GFP wt 19-22 34-36 50-54 D188A FYRN

pMX AID-GFP

b

AID Flag-AID NT

anti-AID

25 —

32 —
25 —

anti-PCNA

AID-GFP GFP-AID NT

62 —
47 —

anti-AID

32 —
25 —

anti-PCNA

c

Colonies/ 10⁹ viable cells

1800
1400
1000
600
200
0

AID AID D187A/D188A Vector

382 429 45

10⁵
10⁴
10³
10²
10

AID AID-A2 19-22 AID-A2 34-36 AID-A2 50-54 Vector

1086 596 17555 90 117

d

Relative GFP MFI

1
0.8
0.6
0.4
0.2
0

▲ GFP + lept B
□ AID-GFP
■ AID DD-GFP
○ AID-GFP + lept B
◉ AID DD-GFP + lept B

Time (h)
1 2 4 6 post CHX

e

Ramos | | | | | Ramos AID^lo

Steady state | Steady state | 2 h Lept B | 5 h Lept B | 5 h Lept B + MG132 | Steady state

No 1ry Ab | | anti-AID

Ramos AID^lo

25 —

anti-AID

32 —
25 —

anti-PCNA

Figure 7
(Di Noia)