THE EMBO JOURNAL

Non-canonical role of Phf5a in DNA repair and antibody switch recombination relies on histone H2A variants regulation

Nasim Begum, Farazul Haque, Andre Stanlie, Afzal Husain, Samiran Mondal, Mikiyo Nakata, Takako Taniguchi, Hisaaki Taniguchi, and Tasuku Honjo,
DOI: 10.15252/embj.2020106393

Corresponding author(s): Tasuku Honjo (honjo@mfour.med.kyoto-u.ac.jp)

Review Timeline:

Submission Date:	31st Jul 20
Editorial Decision:	18th Aug 20
Revision Received:	15th Dec 20
Editorial Decision:	8th Jan 21
Revision Received:	15th Mar 21
Accepted:	22nd Mar 21

Editor: Karin Dumstrei

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

1st Editorial Decision 18th Aug 2020

Dear Tasuku,

Thank you for submitting your manuscript to The EMBO Journal. Your study has now been seen by two referees and their comments provided below.

As you can see from the comments, the referees find the analysis describing a role for Phf5a in DNA repair and class switch recombination (CSR) interesting and both are supportive of publication in The EMBO Journal. They raise a number of points that would be good to resolve in a revised version. This includes extending some of the findings to other cells that the CH12 cell line and ruling out a role for Ph5a splicing function in CSR. Should you be able to extend the findings along the lines suggested by the referees then I would like to invite you to submit a revised manuscript to the EMBO Journal.

I am happy to discuss the raised points further and maybe it would be most helpful to do so via email or a video call.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website:
https://www.embopress.org/page/journal/14602075/authorguide#transparentprocess

We generally allow three months as standard revision time. As a matter of policy, competing manuscripts published during this period will not negatively impact on our assessment of the conceptual advance presented by your study. However, we request that you contact the editor as soon as possible upon publication of any related work, to discuss how to proceed. Should you foresee a problem in meeting this three-month deadline, please let me know in advance and I can grant an extension.

Thank you for the opportunity to consider your work for publication. I look forward to discussing your revisions further.

with best wishes

Karin

Karin Dumstrei, PhD
Senior Editor
The EMBO Journal

Instructions for preparing your revised manuscript:

Please make sure you upload a letter of response to the referees' comments together with the revised manuscript.

Please also check that the title and abstract of the manuscript are brief, yet explicit, even to non-specialists.

When assembling figures, please refer to our figure preparation guideline in order to ensure proper formatting and readability in print as well as on screen:
http://bit.ly/EMBOPressFigurePreparationGuideline

IMPORTANT: When you send the revision we will require
- a point-by-point response to the referees' comments, with a detailed description of the changes made (as a word file).
- a word file of the manuscript text.
- individual production quality figure files (one file per figure)
- a complete author checklist, which you can download from our author guidelines (https://www.embopress.org/page/journal/14602075/authorguide).
- Expanded View files (replacing Supplementary Information)
Please see out instructions to authors
https://www.embopress.org/page/journal/14602075/authorguide#expandedview

Please remember: Digital image enhancement is acceptable practice, as long as it accurately represents the original data and conforms to community standards. If a figure has been subjected to significant electronic manipulation, this must be noted in the figure legend or in the 'Materials and Methods' section. The editors reserve the right to request original versions of figures and the original images that were used to assemble the figure.

Further information is available in our Guide For Authors:
https://www.embopress.org/page/journal/14602075/authorguide

The revision must be submitted online within 90 days; please click on the link below to submit the revision online before 16th Nov 2020.

https://emboj.msubmit.net/cgi-bin/main.plex

Referee #1:

This report examines the role of Phf5a in DNA repair and class switch recombination. A screen of PHD containing proteins by the authors suggested a role of this protein in CSR, and indeed, siRNA knockdown confirmed this result. Phf5a is a small single PHD containing protein that is not well characterized. They go on to show that Phf5a promotes CSR and DNA repair by binding and recruiting p400 to DSBs, thereby promoting H2A variant histone deposition these sites that ultimately facilitate NHEJ. This report is thorough, well controlled with solid data that supports the conclusions, and is well written. While the work could have included KOs of Phf5a in CH12 cells, their siRNA analysis was convincing. I don't have much more to add except the minor comments below, but this paper should be accepted with minor modifications.

Minor:
-Figures 4F-I are discussed at a later part in the manuscript. These data are related to Fig 5 and are discussed after this figure so it could be incorporated into this Figure or as a supplemental Figure.

Also, Fig 4F Y axis number cropped.
-The authors could speculate in the discussion if and how Phf5a is recruited to DSB sites. They could also speculate how it might regulate the SAGA complex, which has also been found to promote CSR, as Phf5a binds to members of this complex.
-Some minor typos throughout that need correcting.

Referee #2:

This work used a candidate screening approach to test the involvement of proteins with PHD-domains (H3K4me3 readers) in class switch recombination in the CH12 B cell line. They identify Phf5a, a splicing factor, as important both for CSR and AID-initiated IgH-cMyc translocations. Phf5a was dispensable for SHM (with caveats) and DNA breakage at the Smu, as shown by convincing assays. Reporter assays in a lung cancer cell line indicated that Phf5a is required for NHEJ but not HR repair of ISceI-induced DNA DSBs. Accordingly, the residual S-junctions in Phf5a-low CH12 showed more microhomology and large insertions, and less recruitment of Ku80. The balance of end resecting enzymes was also altered, with less ExoI and Mre11, but unaffected CtIP. A defect in the early stages of the DDR was evidenced by reduced occupancy of several H2A variants associated to DNA repair. In particular, H2A.Z depletion replicated the NHEJ defect of Phf5a depletion in reporter cell lines. Depletion of p400, the histone chaperone that deposits H2A.Z, also phenocopied most of the defects of Phf5a depletion (CSR, translocations, DNA repair, S region occupancy of DNA repair factors and histones). Since Phf5a depletion reduced p400 occupancy at the Ig but not the vice versa. A structure-function analysis of Phf5a identified critical residues in specific Zn fingers for CSR. AP-MS analysis of Phf5a interactors revealed splicing, transcription and chromatin associated proteins including p400. A good correlation between the ability of some hits to interact with CSR-proficient and -deficient Phf5a mutants and being necessary for CSR (siRNA in CH12) provided beautiful validation of the data.

The data nicely implicates a novel Phf5a > p400 > H2AZ > NHEJ axis in CSR, demonstrating a previously unknown role for Phf5a in chromatin regulation to promote NHEJ. I am overall positive about this work, but a few additional points should be addressed.

1) A major concern is that all the work was done in the CH12 cell line, and there have been previous cases in which factors implicated in CSR in CH12 cells resulted dispensable in primary cells (notably the splicing factor PTBP2). It would be important to report the expression of Phf5a in activated and germinal center B cells compared to naïve B cells; and to at least confirm by knockdown that Phf5a is required for CSR in primary B cells.

2) The Phf5a > p400 > H2AZ > NHEJ axis is quite well supported but two additional controls should be done. First, show that the GLT and AID expression are not affected in the experiments using p400 knockdown in CH12 cells. Second, the bottom panel in fig 7E is not useful (p400 deposits H2A.Z, complementing sip400 cells with Phf5a was not expected to do anything). They should instead show that complementation of the siPhf5a CH12 cells with the Phf5a C72/C75 mutant does not recover H2A.Z occupancy, while the C72 mutant does.

3) This work focuses on the role of Phf5a in DNA repair by recruiting / stabilizing p400 to modify the chromatin environment after DNA damage. But Phf5a has several roles, including in splicing. A role of the Ph5a splicing function in CSR cannot be ruled out with the data provided.
First, the claim that splicing is not affected in Ph5a-low cells is mentioned but has no data to

support it. Splicing analysis would require RNA-seq but I did not see that data. Figure S18 is not useful, the splicing inhibitor used just reduces the transcript level of many factors required for CSR. So, how was splicing checked? The Igh splicing is important for CSR, was it analyzed?
Second, multiple splicing factors associated to Phf5a are also necessary for CSR, as they shown. Do the authors think this is an independent function? Is there any change in p400 at the Igh in cells depleted of other Sf3b factors? An outstanding and very interesting question is how Phf5a recruits p400? Would it work alone or with the spliceosome? Does knocking down other splicing factors (i.e. the Sf3b members found to interact with Phf5a and/or reduce CSR) also affect p400 recruitment and H2A.Z deposition?

Other points to address

The assay to rule out a role for Phf5a in SHM is not optimal. It relies on an overexpressed hyperactive form of AID that is CSR deficient. If the purpose is to pinpoint the activity of Phf5a to a stage after deamination, it would be more appropriate to sequence the Sμ region in the CH12 cells to show that it is mutated by AID equally than the control.

The authors suggest that the end effect of Phf5a in NHEJ is to prevent excessive resection, but this is not shown. Resection is not measured. I am not sure unchanged Ctip alone would be sufficient for claiming more end resection, as far as I understand Ctip usually works with other nucleases, like the MRN complex rather than alone is not sufficient for resection. And other nucleases involved in end resection are actually reduced. I would temper this conclusion.

The manuscript is very dense, with a lot of data and written for a very specialized audience expert in CSR. Perhaps not every experiment needs to be shown? Some data might be peripheral or unnecessary. At least they should try provide brief introductory explanations for some experiments (e.g. the use of Smarca4, top1, etc knockdown is not obvious, LM-PCR and other specialized techniques are not explained).

They should carefully check figures and figure legends for clarity and consistency, as well as to ensure there is enough information for a general audience to understand the experiments. Some examples:
Figure S1 - some arrows are missing in up/down column.
Fig 3 loading control is b-tubulin or b2m? Fig 3F - the signal is relative according to the y axis label. Relative to what? What was it normalized to?
Please indicate the origin of the cell line H1299dA3-1 in the text.
Table S1 requires further details to understand it.
When describing Fig S8 the authors indicate that Phf5a knockdown does not change gene expression while several of the genes are increased (ung2, lig4, etc).
What is the gene used for normalization in each qPCR in different figures? They seem to have not always used the same one.
Fig. S15; it is difficult to understand what is being compared. Clarify AID wt and KO conditions in the fig and legends.

Manuscript EMBOJ-2020-106393

Reply to Referee -1:

This report examines the role of Phf5a in DNA repair and class switch recombination. A screen of PHD containing proteins by the authors suggested a role of this protein in CSR, and indeed, siRNA knockdown confirmed this result. Phf5a is a small single PHD containing protein that is not well characterized. They go on to show that Phf5a promotes CSR and DNA repair by binding and recruiting p400 to DSBs, thereby promoting H2A variant histone deposition these sites that ultimately facilitate NHEJ. This report is thorough, well controlled with solid data that supports the conclusions, and is well written. While the work could have included KOs of Phf5a in CH12 cells, their siRNA analysis was convincing. I don't have much more to add except the minor comments below, but this paper should be accepted with minor modifications.

Minor:

1. Figures 4F-I are discussed at a later part in the manuscript. These data are related to Fig 5 and are discussed after this figure so it could be incorporated into this Figure or as a supplemental Figure. Also, Fig 4F Y axis number cropped.

As the reviewer suggested, we transferred Fig.4F-I to supplementary (Fig.S4F-I).
We have also corrected the Y-axis. Thank you for noticing the error.

2. The authors could speculate in the discussion if and how Phf5a is recruited to DSB sites. They could also speculate how it might regulate the SAGA complex, which has also been found to promote CSR, as Phf5a binds to members of this complex.

- If and how Phf5a is recruited to DSB sites-

To understand the correlation between AID-induced DSBs and Phf5a-mediated H2A.Z regulation, we used AIDER-expressing CH12 cell line where AID can be activated by Tamoxifen (OHT) treatment. Accumulation of γH2AX, Ku80, and Exo1 were evident in the S region in AID-dependent manner (Fig.S13A). But the deposition of H2A.Z and other H2A variants remained unaffected upon AIDER activation. Phf5a knockdown also equally reduced H2A.Z and p400 in AID-deficient CH12 cells (Fig.S13B). Therefore, we speculate a steady state of pre-existing Phf5a/p400 complex at the S region, which is likely required to maintain a well-balanced H2A variant, which in turn is involved in the initiation of DDR signal upon DSB encounter (Discussion, p16 last paragraph).

We also showed that p400 KD does not affect Phf5a recruitment at the S-region. Since Phf5a is a PHD motif containing protein, it may directly bind the S region chromatin through H3K4me3. However, it may not be trivial to distinguish Phf5a-H3K4me3 interaction from Phf5a-U2SnRNP complex association, especially since SF3b1/3 directly interacts with histones. Future works that aim to identify a separation-of-function mutant of Phf5a that completely dissociates from U2SnRNP while still retain binding to the S-region may help to reveal the mechanism of Phf5a recruitment to the S-region. It is worth mentioning that in the

glycerol gradient sedimentation assay (unpublished), Phf5a distribution can be grouped into 3 distinct zones along with chromatin, DNA repair and spliceosomal proteins. Future analysis of chromatin and DNA repair-associated fractions from CSR-stimulated cells may also help in addressing the question.

- How Phf5a might regulate the SAGA complex, which has also been found to promote CSR, as it binds to members of this complex-

The DUB module (*1*) of SAGA has been reported to play a critical role in the DNA repair phase of CSR by deubiquitinating H2BK120 (*2, 3*) . In the absence of DUB components or SAGA deubiquitinase activity, increased H2BK120 ubiquitination was observed, which subsequently impaired H2AX phosphorylation and interfered with the signaling of early damage response.
As Phf5a interacts and /or stabilizes large protein complexes (U2SnRNP, Paf1/Ski8, p400/H2AZ), it can therefore be envisaged that Phf5a may exert a similar effect on SAGA, likely by being associated with SF3B3 in the "Splicing Module" (*1*). While the exact function of this module is yet unknown, should SAGA-Phf5a interaction be proven critical for the function of the DUB module, then loss of Phf5a would increase H2BK120ub and lead to CSR impairment. Similarly, if the HAT module (*1*) is affected, the activity of GCN5/PACF acetyltransferase will be perturbed, which could in turn lead to a reduction in H3K9acS10p and subsequently impacting the recruitment of 14-3-3 and AID to the S region (*4, 5*).

We briefly mentioned the hypothetical possibility in the revised manuscript (Discussion, p17 last para, text color coded in green), and we also willing to revise it further as per reviewer's advice.

[3] Some minor typos throughout that need correcting.
Thank you for noticing. We have carefully checked the text and corrected the typos.

1. D. Helmlinger, L. Tora, Sharing the SAGA. *Trends Biochem Sci* **42**, 850-861 (2017).
2. C. Li *et al.*, The H2B deubiquitinase Usp22 promotes antibody class switch recombination by facilitating non-homologous end joining. *Nat Commun* **9**, 1006 (2018).
3. S. Ramachandran *et al.*, The SAGA Deubiquitination Module Promotes DNA Repair and Class Switch Recombination through ATM and DNAPK-Mediated gammaH2AX Formation. *Cell Rep* **15**, 1554-1565 (2016).
4. G. Li *et al.*, Combinatorial H3K9acS10ph histone modification in IgH locus S regions targets 14-3-3 adaptors and AID to specify antibody class-switch DNA recombination. *Cell Rep* **5**, 702-714 (2013).
5. B. Vaidyanathan, W. F. Yen, J. N. Pucella, J. Chaudhuri, AIDing Chromatin and Transcription-Coupled Orchestration of Immunoglobulin Class-Switch Recombination. *Front Immunol* **5**, 120 (2014).

Reply to Referee-2:

This work used a candidate screening approach to test the involvement of proteins with PHD-domains (H3K4me3 readers) in class switch recombination in the CH12 B cell line. They identify Phf5a, a splicing factor, as important both for CSR and AID-initiated IgH-cMyc translocations. Phf5a was dispensable for SHM (with caveats) and DNA breakage at the Smu, as shown by convincing assays. Reporter assays in a lung cancer cell line indicated that Phf5a is required for NHEJ but not HR repair of ISceI-induced DNA DSBs. Accordingly, the residual S-junctions in Phf5a-low CH12 showed more microhomology and large insertions, and less recruitment of Ku80. The balance of end resecting enzymes was also altered, with less ExoI and Mre11, but unaffected CtIP. A defect in the early stages of the DDR was evidenced by reduced occupancy of several H2A variants associated to DNA repair. In particular, H2A.Z depletion replicated the NHEJ defect of Phf5a depletion in reporter cell lines. Depletion of p400, the histone chaperone that deposits H2A.Z, also phenocopied most of the defects of Phf5a depletion (CSR, translocations, DNA repair, S region occupancy of DNA repair factors and histones). Since Phf5a depletion reduced p400 occupancy at the Ig but not the vice versa. A structure-function analysis of Phf5a identified critical residues in specific Zn fingers for CSR. AP-MS analysis of Phf5a interactors revealed splicing, transcription and chromatin associated proteins including p400. A good correlation between the ability of some hits to interact with CSR-proficient and -deficient Phf5a mutants and being necessary for CSR (siRNA in CH12) provided beautiful validation of the data.

The data nicely implicates a novel Phf5a > p400 > H2AZ > NHEJ axis in CSR, demonstrating a previously unknown role for Phf5a in chromatin regulation to promote NHEJ. I am overall positive about this work, but a few additional points should be addressed.

- A major concern is that all the work was done in the CH12 cell line, and there have been previous cases in which factors implicated in CSR in CH12 cells resulted dispensable in primary cells (notably the splicing factor PTBP2). It would be important to report the expression of Phf5a in activated and germinal center B cells compared to naïve B cells; and to at least confirm by knockdown that Phf5a is required for CSR in primary B cells.

We examined the expression of Phf5a in naïve versus activated B cells (Fig.S2A) and confirmed that it is not only expressed in activated B cells, but the timing also corresponds to that of AID, confirming its novel role in CSR. Publicly available gene expression dataset (GEO, NCBI) also supports our data which also suggests that Phf5a is indeed expressed in germinal center B cells.
We performed both Phf5a and P400 knockdown in primary B cells, and showed that they are required for Ig isotype switching (Fig.S2B,C and Fig.S7C,D).

- 2) The Phf5a > p400 > H2AZ > NHEJ axis is quite well supported but two additional controls should be done. First, show that the GLT and AID expression are not affected in the experiments using p400 knockdown in CH12 cells. Second, the bottom panel in fig 7E is not useful (p400 deposits H2A.Z, complementing sip400 cells with Phf5a was not expected to do anything). They should instead show that complementation of the

siPhf5a CH12 cells with the Phf5a C72/C75 mutant does not recover H2A.Z occupancy, while the C72 mutant does.

We performed the suggested experiment related to p400 KD in CH12 cells, and showed that the GLT and AID transcript are not affected in the absence of p400 (Fig.S7A,B).

We appreciate reviewer's suggestion to perform complementation assay in Phf5a-depleted CH12 cells. The new data provided (Fig.8) show that C72/C75 mutant is unable to restore CSR as well as H2A.Z loss at the IgH locus. On the other hand, the C75 mutant is capable in restoring both CSR as well as H2A.Z level (Result, p14 last para, text colored in blue).

- 3) This work focuses on the role of Phf5a in DNA repair by recruiting / stabilizing p400 to modify the chromatin environment after DNA damage. But Phf5a has several roles, including in splicing. A role of the Ph5a splicing function in CSR cannot be ruled out with the data provided.
  First, the claim that splicing is not affected in Ph5a-low cells is mentioned but has no data to support it. Splicing analysis would require RNA-seq but I did not see that data. Figure S18 is not useful, the splicing inhibitor used just reduces the transcript level of many factors required for CSR. So, how was splicing checked? The Igh splicing is important for CSR, was it analyzed?

As the reviewer suggested, we examined splicing of IgH transcript, as well as H2A.Z, p400 and AID (Fig.S16A-F). We applied conventional approach of transcript variant analysis along with unspliced /spliced transcript comparison. Recently, this method has also been used to analyze IgH locus splicing (Marchalot *et al*, Frontiers in Immunology, 2020).

In addition, we showed RT-PCR amplification of full coding region (CDS) of more than 26 genes, including essential and major repair genes associated with CSR [Fig.S15A (Fig.S7, previously). Primers were specially designed to encompass ATG to STOP codon of each gene, and the image of the entire gel length was shown to assure that there is no difference in the PCR product between siControl and siPhf5a treated samples. The CDS of these genes were ~0.3 kb -9.0 kb length (indicated on the top of each gel panel) and derived from multi-exon-intron containing premRNAs. Without the appropriate splicing, the sizes of these CDS amplification from the respective mRNA would be different. We also provided RT-qPCR data for a quantitative reflection on transcription, where we were required to design separate primer pairs to amplify a short region within the transcript [Fig.S15B (Fig.S8, previously)]. We did not observe any down-regulation of the analyzed transcripts, instead some such as UNG2, showed an elevated expression as the reviewer noticed.

In this context, it is indeed compelling to note that the treatment of splicing inhibitor [Fig.S17 (Fig.S18, previously)] led to an undetectable level of not only UNG2 but also some other transcripts that were not inhibited by Phf5a depletion. A similar observation was noted in the case of Sf3b1 depletion (Fig.S16F). Since this data highlights the distinguishing features between cells treated with splicing inhibitor vs. siPhf5a, we hope that the reviewer will agree with us in including the data in the manuscript. We have discussed the finding in the discussion (p17 top, text colored in blue).

A separate study is underway to understand the difference between Phf5a deficiency, and Sf3b1 deficiency or splicing inhibitor treatment. This study is based on an RNA-seq approach to address the global transcription and splicing profile, which will provide an in-depth view and further complement the deficiency, if any, of our current study.

- Second, multiple splicing factors associated to Phf5a are also necessary for CSR, as they shown. Do the authors think this is an independent function? Is there any change in p400 at the Igh in cells depleted of other Sf3b factors? An outstanding and very interesting question is how Phf5a recruits p400? Would it work alone or with the spliceosome? Does knocking down other splicing factors (i.e. the Sf3b members found to interact with Phf5a and/or reduce CSR) also affect p400 recruitment and H2A.Z deposition?

The H2A.Z deposition was not affected by knocking down Sf3b3 or Sf3b14a (not shown). We were unable to generate any comparable dataset for the key splicing factor Sf3b1 as its KD caused proliferation arrest and drastically reduces AID and p400 expression (Fig.S14). Since there are many U2 SnRNP subunits, it is difficult to completely rule out their involvement in our current scope of investigation. Studies involving key splicing factors dampens our enthusiasm as their KD often affects key genes related to CSR, particularly AID.

It is indeed an outstanding question as to how Phf5a recruits p400. We currently speculate the involvement of multiple factors as one possible mechanism; necessitating an extensive study from p400 perspective. Moreover, the concept of phase separation nuclear compartments is gaining traction to explain complex functions including DNA repair and splicing (Li et al, 2020, Kilic et al, 2019). This is an intriguing possibility to be explored in future.

References

Kilic S, Lezaja A, Gatti M, Bianco E, Michelena J, Imhof R, Altmeyer M (2019) Phase separation of 53BP1 determines liquid-like behavior of DNA repair compartments. EMBO J 38: e101379

Li W, Hu J, Shi B, Palomba F, Digman MA, Gratton E, Jiang H (2020) Biophysical properties of AKAP95 protein condensates regulate splicing and tumorigenesis. Nat Cell Biol 22: 960-972

Other points to address

- (a) The assay to rule out a role for Phf5a in SHM is not optimal. It relies on an overexpressed hyperactive form of AID that is CSR deficient. If the purpose is to pinpoint the activity of Phf5a to a stage after deamination, it would be more appropriate to sequence the Sμ region in the CH12 cells to show that it is mutated by AID equally than the control.

We have added the SHM data based on the Sμ sequencing results in the Supplementary (Fig.S3C,D).

Variable (V) region SHM occurs in BL2, but not in CH12 cells. On the other hand, CSR occurs in CH12 but not in BL2. Therefore, BL2 and CH12 cells are considered model cell lines to study V-SHM and CSR, respectively. Since we specifically wanted to address V region SHM, we selected BL2 line.

- (b) The authors suggest that the end effect of Phf5a in NHEJ is to prevent excessive resection, but this is not shown. Resection is not measured. I am not sure unchanged Ctip alone would be sufficient for claiming more end resection, as far as I understand Ctip usually works with other nucleases, like the MRN complex rather than alone is not sufficient for resection. And other nucleases involved in end resection are actually reduced. I would temper this conclusion.

Our interpretation on break end "resection" was mainly based on the strong impairment observed on NHEJ and an increase in microhomology and insertion at the recombination junctions (AID- and I-SceI-induced DNA breaks) upon Phf5a depletion. These phenomena are associated with DSB resection, which exposes short/long single stranded DNA, leading to various types of end processing not conducive to NHEJ. A site-specific DNA end resection data also supports this conclusion (Fig.S4D,E). This was further bolstered by the fact that the occupancy of DNA repair factors such as Ku80 – known to prevent resection (REF) – in IgH locus are also decreased in Phf5a depleted cells (Fig.5C).

Since Mre11 was not completely depleted from the break site, CtIP may promote the activity of MRN complex. Moreover, CtIP can work as a complex with other endonucleases/helicases and may process a subset of DSB ends through its endonuclease activity. We have edited the text with necessary references to clarify our previous statement (Result, p8-9, last para, text colored in blue).

- (c) The manuscript is very dense, with a lot of data and written for a very specialized audience expert in CSR. Perhaps not every experiment needs to be shown? Some data might be peripheral or unnecessary. At least they should try provide brief introductory explanations for some experiments (e.g. the use of Smarca4, top1, etc knockdown is not obvious, LM-PCR and other specialized techniques are not explained). They should carefully check figures and figure legends for clarity and consistency, as well as to ensure there is enough information for a general audience to understand the experiments. Some examples:

As reviewer suggested, we removed some of the Supplementary data and provided additional explanations and appropriate schemes (e.g. Smarca4 and Top1, and LMPCR)

- Figure S1 - some arrows are missing in up/down column.

We have added the arrows to indicate up/down.

- Fig 3 loading control is b-tubulin or b2m? Fig 3F - the signal is relative according to the y axis label. Relative to what? What was it normalized to?

Thank you for pointing out. We noticed that the control panel was mislabeled, which should be Actin. The y axis is now labeled as "Relative signal (% Input)" so that the DNA break level can be evaluated in CIT (+) samples relative to CIT (-). AID is expressed only after CIT treatment and DNA breaks are subsequently initiated in the IgH locus. This CIT-dependent, AID-induced DNA break was not observed in the b-actin locus. To further clarify this point, the schematics of the 2 types of DNA break assays employed in this study – LMPCR (Fig.3D,E) and End Labeling (Fig.3F) – are illustrated in Supplementary (Fig.S3E,F).

- Please indicate the origin of the cell line H1299dA3-1 in the text.

The detail of this NHEJ reporter cell line was previously described by Ogiwara et al which we referred and also mentioned in the Method section of the manuscript. We received the line from these authors.

- Table S1 requires further details to understand it.

The table provides the information on major Phf5a associated proteins identified by Mass-spectrometry (MS). We added an explanation in the legend.

- When describing Fig S8 the authors indicate that Phf5a knockdown does not change gene expression while several of the genes are increased (ung2, lig4, etc).

We intended to say that the expression was not negatively affected and we have carefully stated this statement in the revised manuscript.

- What is the gene used for normalization in each qPCR in different figures? They seem to have not always used the same one.

We used either Hprt/Tubulin for RT-qPCR normalization, which was indicated in the Y-axis or in the legend, and also in the Method.

- Fig. S15; it is difficult to understand what is being compared. Clarify AID wt and KO conditions in the fig and legends.

In the revised manuscript, we provided illustration and also improved text in the legend (Fig.S13).

1st Revision - Editorial Decision 8th Jan 2021

Dear Prof. Honjo,

Thank you for submitting your revised manuscript to The EMBO Journal. Your study has now been seen by the original referee #2 whose comments are provided below. As you can see the referee appreciates the added changes and supports publication here. However, the referee also notes that the M&M section needs to be updated to better describe the techniques used. Will you please carefully go through that section and make sure that it is a detailed as possible. Also, I see that the entire experimental methods (please re-label that section as Materials and Methods) is in the appendix. It is OK to have part of the M&M in the appendix, but the main part should be in the main MS file.

When you submit a revised version will you also take care of the following points

- In some of the gels it appears that there are some unmarked splices between the lanes (like Figure 2B, S8G and S16B) will you please carefully look at the blots. Can you also make sure that the white line used to mark the splice is thick enough to see. Please also provide source data for all the gels (one PDF file per figure). The source data should contain the original, uncropped and unprocessed scans of the gels used in the figure. The PDF files should be labeled with the appropriate figure/panel number, and should have molecular weight markers; further annotation could be useful but is not essential. The PDF files will be published online with the article as supplementary "Source Data" files.

- You are missing a Data Availability section. This is the place to enter accession numbers etc. As far as I can see no data is generated that needs to be deposited in a database. If this is correct please state: This study includes no data deposited in external repositories. Please place it after the Materials and methods and before Acknowledgements

- You can only have 5 keywords, you have at the moment 6

- COI section is missing

- In the acknowledgement section NM should be MN. Please double check this

- The text is missing callouts to Appendix Tables S2-S5

- For all Appendix figures and tables please add Appendix to the file name

- We include a synopsis of the paper (see http://emboj.embopress.org/). Please provide me with a general summary statement and 3-5 bullet points that capture the key findings of the paper.

- We also need a summary figure for the synopsis. The size should be 550 wide by [200-400] high (pixels). You can also use something from the figures if that is easier.

- I have asked our publisher to do their pre-publication checks on the paper. They will send me the file within the next few days. Please wait to upload the revised version until you have received their comments.

- Please correct Summary to Abstract

That should be all. Let me know if you have any further questions.

With best wishes

Karin

Karin Dumstrei, PhD
Senior Editor
The EMBO Journal

Instructions for preparing your revised manuscript:

Please check that the title and abstract of the manuscript are brief, yet explicit, even to non-specialists.

When assembling figures, please refer to our figure preparation guideline in order to ensure proper formatting and readability in print as well as on screen:
https://bit.ly/EMBOPressFigurePreparationGuideline

IMPORTANT: When you send the revision we will require
- a point-by-point response to the referees' comments, with a detailed description of the changes made (as a word file).
- a word file of the manuscript text.
- individual production quality figure files (one file per figure)
- a complete author checklist, which you can download from our author guidelines (https://www.embopress.org/page/journal/14602075/authorguide).
- Expanded View files (replacing Supplementary Information)
Please see out instructions to authors
https://www.embopress.org/page/journal/14602075/authorguide#expandedview

Please remember: Digital image enhancement is acceptable practice, as long as it accurately represents the original data and conforms to community standards. If a figure has been subjected to significant electronic manipulation, this must be noted in the figure legend or in the 'Materials and Methods' section. The editors reserve the right to request original versions of figures and the original images that were used to assemble the figure.

Further information is available in our Guide For Authors:
https://www.embopress.org/page/journal/14602075/authorguide

The revision must be submitted online within 90 days; please click on the link below to submit the revision online before 8th Apr 2021.

https://emboj.msubmit.net/cgi-bin/main.plex

Referee #2:

The authors have answered all my concerns and provided relevant additional data. I believe the data supports their conclusions. I have no further concerns and support publication without any further modifications.

However, I would advise that they check the methods section, I believe some of the techniques added in the revision are not described there (methods for primary B cells and some new assays like the RED based end resection at IgH).

REF: EMBOJ2020-106393R

Response to Referee #2:

The authors have answered all my concerns and provided relevant additional data. I believe the data supports their conclusions. I have no further concerns and support publication without any further modifications.

However, I would advise that they check the methods section, I believe some of the techniques added in the revision are not described there (methods for primary B cells and some new assays like the RED based end resection at IgH).

Thank you for noticing the missing sections under Materials and Methods. In the revised version, we have now described primary B cell work (p20), Sμ–SHM analysis (p22), and DNA end resection assay (p26). Relevant reagents were updated in the Appendix Tables S2 and S3. Primers related to splicing and/or transcription are listed in Appendix Table S6.

2nd Revision - Editorial Decision 22nd Mar 2021

Dear Tasuku,

Thank you for submitting your revised manuscript to The EMBO Journal. I have now had a chance to take a look at everything and I appreciate the introduced changes.

I am therefore very pleased to accept the manuscript for publication here.

Congratulations on a nice study!

with best wishes Karin

Karin Dumstrei, PhD
Senior Editor
The EMBO Journal

Please note that it is EMBO Journal policy for the transcript of the editorial process (containing referee reports and your response letter) to be published as an online supplement to each paper. If you do NOT want this, you will need to inform the Editorial Office via email immediately. More information is available here: https://emboj.embopress.org/about#Transparent_Process

Your manuscript will be processed for publication in the journal by EMBO Press. Manuscripts in the PDF and electronic editions of The EMBO Journal will be copy edited, and you will be provided with page proofs prior to publication. Please note that supplementary information is not included in the proofs.

Should you be planning a Press Release on your article, please get in contact with embojournal@wiley.com as early as possible, in order to coordinate publication and release dates.

If you have any questions, please do not hesitate to call or email the Editorial Office. Thank you for your contribution to The EMBO Journal.

** Click here to be directed to your login page: https://emboj.msubmit.net

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Tasuku Honjo
Journal Submitted to: EMBO J
Manuscript Number: EMBOJ-2020-106393

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	Power calculation was based on the published literatures, our previous experience, and intial experiemnts.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	Based on our experience, for CSR and routine expression analysis, we use a minimum group size of n=3 . In the present study, the group size was n=3 for gene knckdown, CSR and other subsequent analysis.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	The mice were aged matched and randomly selected irrespective of their sex.
For animal studies, include a statement about randomization even if no randomization was used.	See above.
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	Not blinded
4.b. For animal studies, include a statement about blinding even if no blinding was done	Not blinded
5. For every figure, are statistical tests justified as appropriate?	Performed stastiscal analysis for all the critical data and also indicated in the respective figure.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	The stastiscal analysis performed was also mentioned.

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jij.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

Is there an estimate of variation within each group of data?	NO
Is the variance similar between the groups that are being statistically compared?	NA

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	The list of antibodies along with their catalogue numbers is provided in the Appendix Table S3.
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Cell lines used are described in the method section. We perform routine check up on the mycoplama contamination for cell lines cultured in the lab.

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	C57BL6
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	The experiemnts were carried out as per the guidelines provided by the Kyoto University ethical committee.
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	We confirm full compliance with the NIH/MRC guidelines.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	NA
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	NA
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA