

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ

Institut de Génétique et de Biologie Moléculaire et Cellulaire (UMR_S1258)

THÈSE présentée par :

Vincent HEYER

soutenue le : **03 juillet 2024**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Sciences de la vie et de la santé

**Étude de l'impact des facteurs induits par l'hypoxie
sur l'expression de l'enzyme AID au cours du
processus de commutation isotypique des
immunoglobulines**

Thèse dirigée par :

Dr. Frédéric GROS MCF, HDR, CRBS, Strasbourg

Rapporteurs:

Dr. Rima HADDAD MCF, HDR, Paris Saclay

Dr. Sandrine ROULLAND DR, HDR, CIML, Marseille

Examineurs

Pr. Céline CLEMENT PU, INSPE, Strasbourg

Pr. Fabrice FAVRET PU, Faculté des sciences du sport, Strasbourg

Dr. Morgane PERDOMINI Directeur R&D, Flamingo Th., Strasbourg

Mme Carole MAILLER Service de Formation Continue, Strasbourg

Table des matières

| | |
|--|-----------|
| Chapitre 1 - Introduction | 9 |
| 1 - Préambule | 9 |
| 2 - Les origines de l'immunologie | 11 |
| 1 - Les épidémies | 12 |
| 2 - De la vaccine à la rage | 14 |
| 3 - La réponse humorale | 17 |
| 1 - La réaction inflammatoire | 17 |
| 2 - L'immunité innée et adaptative | 17 |
| 3 - Les cellules dendritiques | 18 |
| 4 - Les centres germinatifs | 19 |
| 5 - Les cellules T | 19 |
| 6 - L'activation des lymphocytes B | 20 |
| 7 - Les immunoglobulines | 21 |
| 4 - Les mécanismes de diversification des Immunoglobulines | 23 |
| 5 - Activation-Induced cytidine Deaminase (AID) | 27 |
| 1 - Découverte et fonction | 27 |
| 2 - Structure | 27 |
| 3 - AID et le cancer | 28 |
| 4 - Le modèle de déamination de l'ADN | 29 |
| 5 - Les complexes protéiques régulateurs de AID | 30 |
| 6 - Les cellules B et l'hypoxie | 33 |
| 7 - Le complexe HIF | 35 |
| 1 - Structure des HIFs | 35 |
| 2 - Régulation des HIFs | 36 |
| 8 - Objectif de ma thèse | 39 |
| Chapitre 2 - Résultats | 41 |
| 1 - Identification des co-facteurs d'AID | 41 |
| 1 - Extraction des protéines cytoplasmiques et nucléaires | 42 |
| 2 - Immuno-précipitation en tandem | 42 |
| 3 - Découverte du complexe KAP1, HP1 et AID | 43 |
| 4 - Caractérisation des complexes protéiques associés à AID par MudPIT | 45 |
| 5 - Co-immunoprécipitation des partenaires de AID | 47 |
| 6 - Le complexe Polymerase Associated Factor (PAF) | 48 |
| 7 - Le complexe Cohésine | 48 |
| 8 - Le Mediator | 49 |
| 2 - Le Système d'Expression Universel (UES) | 51 |
| 1 - Introduction | 51 |
| 2 - Principe | 52 |

| | |
|--|-----------|
| 3 - Les vecteurs de destination | 53 |
| 4 - Les modules | 54 |
| 5 - Le système CRISPR/Cas9 | 55 |
| 6 - Développement du système UES | 57 |
| 7 - Un criblage à l'échelle du génome..... | 59 |
| 8 - Perspectives | 60 |
| 3 - Étude du complexe protéique HIF..... | 63 |
| 1 - Hif1a est nécessaire pour une CI optimal dans les cellules CH12 | 63 |
| 2 - L'induction de l'ARNm AID est compromise dans les cellules <i>Hif1a</i> ^{-/-} | 67 |
| 3 - Hif1b est nécessaire pour l'expression optimale d'AID et de l'efficacité de la CI | 67 |
| 4 - Le défaut de CI dans les cellules B déficientes en Hif1a ou en Hif1b est dû à l'expression sous-optimale d'AID | 70 |
| Chapitre 3 - Discussion..... | 73 |
| 1 - Identification des co-facteurs de AID | 73 |
| 2 - Impact de HIF sur l'expression de AID et de la CI | 74 |
| Chapitre 4 - Perspectives..... | 79 |
| Manuscrit | 81 |
| Graphical Abstract..... | 91 |
| Matériel et Méthodes..... | 93 |
| Références | 97 |
| Annexes | 101 |

Liste des figures

| | |
|--|----|
| Figure 1: La peste noire. | 13 |
| Figure 2: Pasteur et la Théorie des Germes. | 15 |
| Figure 3: Activation des lymphocytes B. | 21 |
| Figure 4: Structure du récepteur des cellules B. | 23 |
| Figure 5: Mécanismes de génération de la diversité des anticorps. | 24 |
| Figure 6: L'hypermutation somatique. | 25 |
| Figure 7: La commutation isotypique. | 25 |
| Figure 8: Activation induced cytidine deaminase (AID)..... | 28 |
| Figure 9: AID et cancer. | 29 |
| Figure 10: Le modèle de désamination de l'ADN..... | 30 |
| Figure 11: Tensions et gradients d'oxygène rencontrés au cours de la vie d'une cellule B..... | 34 |
| Figure 12: Structure des facteurs HIFs. | 36 |
| Figure 13: Mécanismes de détection de l'oxygène par les cellules et stimulation par les cytokines dans les cellules B (CH12). | 38 |
| Figure 14: Création de la lignée CH12 sur exprimant AID. | 42 |
| Figure 15: Identification des partenaires d'AID. | 43 |
| Figure 16: Chromatographie sur colonne de sépharose..... | 45 |
| Figure 17: Approches GBA et MudPIT..... | 46 |
| Figure 18: Bilan des protéines identifiées par spectrométrie de masse..... | 47 |
| Figure 19: Co-immunoprécipitations des différents complexes associés avec AID...48 | |
| Figure 20: Digestion et ligation simultanée des fragments d'ADN..... | 52 |
| Figure 21: Vue d'ensemble de la procédure d'assemblage du vecteur d'expression. | 53 |
| Figure 22: Les vecteurs de destinations. | 54 |

| | |
|---|----|
| Figure 23: Liste non exhaustive des modules disponibles..... | 55 |
| Figure 24: Le système CRISPR-Cas9. | 56 |
| Figure 25: Création d'un vecteur CRISPR-Cas9 avec deux guides ARN..... | 56 |
| Figure 26: Knock-in par microhomologie (MMEJ)..... | 57 |
| Figure 27: La dCas9..... | 58 |
| Figure 28: Visualisation de l'hétérochromatine avec CRISPR..... | 58 |
| Figure 29: UES. Une gamme non exhaustive..... | 59 |
| Figure 30: Le crible CRISPR/Cas9 à l'échelle du génome..... | 60 |
| Figure 31: Élaboration du knock-out du gène Hif1a..... | 64 |
| Figure 32: Détermination de l'efficacité du knock-out du gène Hif1a..... | 64 |
| Figure 33: Marquage des cellules stimulées au CFSE. | 65 |
| Figure 34: Infection rétrovirale avec pMX-PIE. | 65 |
| Figure 35: Western Blot de la reconstitution de HIF-1 α | 66 |
| Figure 36: Restauration du phénotype Hif1a ^{-/-} (FACS). | 66 |
| Figure 37: Restauration du phénotype Hif1a ^{-/-} (ARN, protéine). | 67 |
| Figure 38: Élaboration du knock-out du gène Hif1b..... | 68 |
| Figure 39: Marquage des cellules stimulées au CFSE. | 68 |
| Figure 40: Restauration du phénotype Hif1b ^{-/-} (FACS). | 69 |
| Figure 41: Restauration du phénotype Hif1b ^{-/-} (ARN, protéine). | 69 |
| Figure 42: Le défaut de CI dans les cellules B déficientes en Hif1a ou en Hif1b est dû à l'expression sous-optimale d'AID. | 71 |
| Figure 43: Compensation de l'expression de la protéine AID..... | 77 |

Liste des abréviations

ADN : Acide désoxyribonucléique
ADNc : Acide désoxyribonucléique complémentaire
AID : Activation-Induced cytidine Deaminase
APE : endonucléase apurinique/apyrimidique
APOBECs : mRNA editing enzyme catalytic polypeptide-like
ARN : acide ribonucléique
ARNm : acide ribonucléique messenger
BC100 : milieu d'immuno-précipitation
BCR : récepteur des lymphocytes B
BER : Base Excision Repair
bHLH : basic-helix loop helix
CAD : transactivation à l'extrémité C-terminale
Cas9 : CRISPR associated protein 9
CD-40L : CD40 Ligand
CFP : cyan fluorescent protein
CFSE : carboxyfluorescéine diacétate succinimidyl ester
CI : commutation isotypique
CMH II : complexe majeur d'histocompatibilité de classe II
CMV : promoteur du cytomégalovirus
CRISPR : Clustered Regularly Interspaced Short Palindromic Repeats
CTR9 : RNA Polymerase-Associated Protein CTR9 Homolog
DAMP : éléments liés à des dommages cellulaires
DSB : cassure double brin de l'ADN
DZ : zone sombre (dark zone)
EPO : Érythropoïétine
FACS : Fluorescence Activated Cell Sorting
FAM72a : Family with Sequence Similarity 72, Member A
FDCs : cellules dendritiques folliculaires
GBA : analyse par découpe de gel
GC : centre germinatif
GFP : Green Fluorescent Protein
GFP : green fluorescent protein
HIF-1 : Facteur 1 Inductible par l'Hypoxie

HIF-1a : Facteur 1 Inductible par l'Hypoxie alpha
HIF-1b : Facteur 1 Inductible par l'Hypoxie bêta ou ARNt
HIF-2a : Facteur 2 Inductible par l'Hypoxie alpha, ou EPAS1
HMS : hypermutation somatique
HP1 : heterochromatin protein 1
HPRT : Hypoxanthine Phosphoribosyltransferase 1
HRE : élément de réponse à l'hypoxie
ID : domaine d'inhibition
Ig : immunoglobuline
IGBMC : Institut de Génétique et de Biologie Moléculaire et cellulaire
IgH : chaine lourde de l'immunoglobuline
IgL : chaine légère de l'immunoglobuline
IL-4 : interleukin 4
IP : immuno-précipitation
KAP1 : KRAB domain-associated protein 1
kDa : kilo Dalton
LC-MS : Spectrométrie de masse par chromatographie en phase liquide
LTR : long terminal repeat sequence
LZ : zone claire (light zone)
MAMP : éléments associés à la présence de microorganismes
MMEJ : jonction d'extrémité alternative non homologue
MMR : mismatch repair
MO : moelle osseuse
MSH2 : MutS protein homolog 2
MSH6 : MutS protein homolog 6
MudPIT : Multidimensionnel Protein Identification Technology
NE : extrait nucléaire
NES : nuclear export signal
NHEJ : Non-Homologous End Joining, jonction non homologue de l'extrémité libre
Nipbl : Nipped-B-like protein
NLS : nuclear localization signal
ODD : oxygen-dependent degradation domain
PAF : Polymerase Associated Factor
PAS : Per-Arnt-Sim
PCR : polymerase chain reaction

pO₂ : pression partielle d'oxygène
PRR : Pattern Recognition Receptors
pVHL : facteur suppresseur de tumeur von Hippel-Lindau
R : région constante de l'immunoglobuline
RH : recombinaison homologue
RT-qPCR : Reverse Transcription – Quantitative PCR
S100 : extrait cytoplasmique
SMC (1/3) : Structural Maintenance of Chromosomes
TAD : domaine de transactivation
TCR : récepteur des lymphocytes T
TGF- β : Transforming growth facteur bêta
UES : Système de Clonage Universel
UNG : uracil-DNA glycosylase
UNG : Uracile DNA Glycosylase
V : région variable de l'immunoglobuline
VEGF : Vascular endothelial Growth factor
Wapl : Wings apart-like protein homolog
WT : wild type
XPO1 : exportine 1
YFP : yellow fluorescent protein

Chapitre 1 - Introduction

1 - Préambule

L'histoire de l'humanité est profondément entrelacée avec la lutte contre les maladies, remontant jusqu'au Néolithique, période marquée par une transition cruciale de modes de vie nomades de chasseurs-cueilleurs à des sociétés agricoles sédentaires.

Alors que les premières communautés humaines ont évolué au sein de la nature, leur mode de vie de chasseur-cueilleur, caractérisé par une faible densité de population et une mobilité constante, a contribué à façonner la dynamique des maladies de manière significative. Au cours de la période paléolithique, les groupes de chasseurs-cueilleurs se déplaçaient en fonction des saisons et des ressources disponibles, ce qui limitait les contacts entre différentes communautés. Cette dispersion géographique réduisait naturellement les opportunités de transmission des maladies, créant ainsi un environnement où la propagation des pathogènes était entravée. De plus, la dépendance des chasseurs-cueilleurs vis-à-vis de la diversité des ressources alimentaires dans leur environnement contribuait probablement à une alimentation équilibrée, minimisant ainsi la prévalence des maladies liées à la malnutrition. Cependant, bien que les sociétés de chasseurs-cueilleurs aient bénéficié de ces facteurs limitant la diffusion des maladies, elles n'étaient pas exemptes de problèmes de santé. Des éléments archéologiques mettent en lumière la manifestation d'infections bactériennes, virales et parasitaires, indiquant ainsi que, même au sein de ces sociétés anciennes, les individus devaient faire face à des enjeux sanitaires.

L'histoire prend un tournant crucial avec le passage à l'agriculture au cours du Néolithique. Cette transition a entraîné l'émergence de communautés sédentaires, caractérisées par une densité de population accrue, la domestication d'animaux et le stockage de surplus alimentaires. Ces changements ont fourni un terrain fertile pour la diffusion des maladies, avec des conditions favorables à la propagation des agents pathogènes.

Bien que les chasseurs-cueilleurs aient pu bénéficier d'une certaine immunité face à la diffusion des maladies en raison de leur mode de vie spécifique, l'histoire de l'humanité est marquée par une constante adaptation aux défis de santé. Les transitions dans les modes de vie ont influencé la manière dont les maladies se sont

propagées, ouvrant la voie à de nouvelles dynamiques épidémiologiques à mesure que l'homme a évolué au fil des siècles.

Une citation saisissante de Thucydide, le célèbre historien de la Grèce antique, offre une description poignante de la maladie. Il affirme que « toute science humaine se révélait impuissante ; malgré les supplications incessantes dans les temples et les consultations des oracles, toutes les tentatives étaient vaines, et finalement, la résignation s'imposa face au fléau ». Cependant, l'auteur souligne que « ceux qui avaient survécu à la maladie témoignaient d'une grande compassion envers les mourants et les malades, car ayant déjà enduré le mal, ils se sentaient en sécurité ; les rechutes ne se révélaient pas fatales ». Ce passage du Ve siècle avant notre ère semble anticiper des concepts qui trouveront écho dans ce que nous appelons aujourd'hui l'immunologie.

2 - Les origines de l'immunologie

Chaque jour, notre organisme fait face à une pléthore de microorganismes, tels que des bactéries, des virus, des parasites et des champignons, certains portant le potentiel de déclencher des maladies. Pour répondre à cette diversité de menaces, l'évolution nous a légué des systèmes de défense de plus en plus sophistiqués, façonnés au cours de millions d'années. Cette coordination complexe de mécanismes protecteurs forme ce que nous connaissons sous le nom de système immunitaire, composé d'une vaste gamme de cellules spécialisées et de molécules dédiées à la reconnaissance, à l'élimination et à la neutralisation des agents pathogènes.

Le rôle du système immunitaire dans la préservation de l'intégrité des tissus est d'une importance capitale. Au-delà de sa fonction principale de défense contre les infections, le système immunitaire contribue activement à maintenir la santé des tissus en orchestrant une surveillance constante. Un aspect particulièrement crucial de cette surveillance est l'immunosurveillance des tumeurs. C'est un mécanisme sophistiqué par lequel le système immunitaire identifie et cible les cellules cancéreuses émergentes. Les cellules cancéreuses, en raison de leurs mutations génétiques et de leur comportement anormal, peuvent parfois échapper à la détection initiale du système immunitaire. Cependant, dans de nombreux cas, le système immunitaire est capable de reconnaître ces cellules défectueuses et de les éliminer avant qu'elles ne se multiplient de manière incontrôlée.

Au fil du temps, le système immunitaire a développé une capacité remarquable d'adaptation pour anticiper et contrer de nouvelles menaces émergentes. Notre corps est ainsi doté d'une véritable armée biologique, prête à détecter et à combattre tout intrus pouvant compromettre notre santé. La complexité et la diversité des composants du système immunitaire témoignent de son évolution constante, soulignant l'ingéniosité de la nature dans la préservation de la vie.

Parallèlement à tous les autres mammifères, notre système immunitaire se divise en deux facettes essentielles : le système immunitaire inné et le système immunitaire adaptatif.

Le système immunitaire inné, considéré comme le socle initial de notre défense, joue un rôle crucial dans notre capacité à repousser efficacement la plupart des corps

étrangers rencontrés quotidiennement. Il agit rapidement en réponse à une menace, formant le premier rempart de notre immunité. Bien que sa réaction soit spécifique, elle se concentre sur la reconnaissance de caractéristiques générales des agents pathogènes, assurant une protection immédiate contre les menaces courantes. De plus, le système immunitaire inné sert de plateforme d'activation pour le système immunitaire adaptatif. Ce dernier, une fois activé, orchestre une réponse plus ciblée et sophistiquée face à des agents pathogènes spécifiques. L'une des caractéristiques remarquables du système immunitaire adaptatif est sa capacité à développer une mémoire immunologique. Cette mémoire permet au corps de reconnaître et de réagir de manière plus rapide et efficace lors d'une exposition ultérieure aux mêmes agents pathogènes, offrant ainsi une protection accrue et une réponse spécifique adaptée à chaque menace.

Lorsque le système immunitaire inné atteint ses limites, c'est le système immunitaire adaptatif qui prend le relais, avec l'implication cruciale des lymphocytes T et B. Les lymphocytes CD4+ coordonnent la réponse immunitaire, tandis que les lymphocytes CD8+ ont la capacité de détruire les cellules renfermant le pathogène ou des cellules cancéreuses. Simultanément, les lymphocytes B produisent après différenciation en plasmocytes, des anticorps capables de reconnaître, neutraliser et contribuer à l'élimination du pathogène. Ces cellules se distinguent par la présence d'un récepteur spécialisé à leur surface, chacun possédant théoriquement une spécificité unique.

En combinant ces spécificités, il devient possible de reconnaître potentiellement n'importe quel agent pathogène avant même qu'il ne soit rencontré. Parallèlement, des lymphocytes T et B de mémoire sont générés, facilitant une réaction plus rapide en cas de réinfection par le même pathogène. Lors d'une infection, les lymphocytes B ont la capacité de diversifier à nouveau les anticorps qu'ils produisent, les rendant plus spécifiques et efficaces. Cela permet de générer une réponse hautement ciblée contre tout pathogène.

1 - Les épidémies

Au cours des deux derniers millénaires, l'humanité a été façonnée par des épidémies dévastatrices, dont la peste d'Antonin, qui a ravagé Rome au IIIe siècle de notre ère, et la peste justinienne, qui a presque anéanti la moitié de la population romaine entre le Ve et le IXe siècle. Ces fléaux ont jeté de sombres nuages sur des civilisations

florissantes. Au XIV^e siècle, la terreur a continué avec la peste noire (figure 1), une pandémie qui a sévi pendant près de trois siècles, laissant des communautés entières décimées.



Figure 1: La peste noire.

La peste noire, également connue sous le nom de peste bubonique, a été l'une des pandémies les plus dévastatrices de l'histoire. Elle a émergé en Asie centrale, probablement en Chine, et s'est propagée en Europe au milieu du XIV^e siècle, vers 1347. La bactérie *Yersinia pestis*, transmise par les puces des rats, était l'agent pathogène responsable de cette maladie. © Claudio Colombo - stock.adobe.com

La multitude de maladies transmissibles qui ont affligé nos ancêtres a maintenu l'espérance de vie à un niveau désespérément bas, stagnante à environ 35-40 ans en moyenne de l'Antiquité jusqu'au milieu du XIX^e siècle. Cependant, un changement spectaculaire s'est opéré grâce à la révolution immunologique, à l'amélioration des conditions d'hygiène et à la découverte des antibiotiques. Ces avancées scientifiques majeures ont contribué à propulser l'espérance de vie à des sommets inégalés, pour atteindre en moyenne 75 ans. Cette révolution a marqué une ère où la compréhension des maladies infectieuses a progressé, permettant de réduire, voire d'éliminer certaines épidémies qui jadis décimaient des populations entières. Actuellement, nous ne succombons généralement plus aux maladies infectieuses qui ont longtemps meurtri notre histoire. Cependant, de nouveaux défis de santé émergent avec la transmission de nouveaux agents pathogènes, notamment ceux capable de franchir les barrières inter-espèces, comme en témoigne la pandémie de Covid-19.

Bien que la révolution immunologique ait considérablement amélioré notre capacité à combattre les infections, elle a également mis en lumière de nouvelles vulnérabilités. En effet, la victoire sur les épidémies a révélé des maladies jusque-là moins

prépondérantes dans un paysage largement dominé par les infections. Les cancers et les maladies cardiovasculaires, autrefois dans l'ombre des épidémies, se révèlent désormais comme des défis majeurs pour la santé publique. Cette évolution souligne la complexité et la constante évolution des enjeux sanitaires, nécessitant une adaptation continue de la science médicale et de la société pour garantir la santé et le bien-être de l'humanité.

2 - De la vaccine à la rage

Au XVIII^e siècle, la variole représentait un fléau redoutable, entraînant la mort d'une personne sur deux après contraction de la maladie. La première tentative de traitement, connue sous le nom de "variolisation", consistait à inoculer du liquide prélevé dans les pustules de patients atteints à des individus sains. Originnaire de Chine, la variolisation a été introduite en Europe au début du XVIII^e siècle. La toute première inoculation a été réalisée à Constantinople en 1701.

Cette pratique reposait sur l'observation selon laquelle les survivants d'une infection ne semblaient pas réinfectés. Cependant, avec un taux de 1 à 2 % des sujets développant effectivement la maladie, cette méthode fut considérée comme dangereuse.

Edward Jenner a joué un rôle crucial à la fin du XVIII^e siècle en introduisant une innovation majeure. Au lieu d'utiliser le liquide provenant de pustules humaines, il opta pour l'inoculation de liquide provenant de pustules de vaches atteintes de la maladie bovine, également appelée "vaccine". Jenner avait en effet observé que les fermières qui traient les vaches semblaient immunisées contre la variole.

Un siècle plus tard, Louis Pasteur développa le concept de "vaccination" et démontra que cette forme d'immunité pouvait s'appliquer à divers microorganismes (Figure 2).



Figure 2: Pasteur et la Théorie des Germes.

Pasteur a contribué de manière significative à la formulation de la théorie des germes, qui stipule que de nombreux processus de fermentation et de maladies sont causés par la présence de microorganismes. Cela a réfuté la croyance antérieure en la génération spontanée, selon laquelle certains organismes pouvaient apparaître sans parenté préexistante.

Il contribua également à la création du célèbre vaccin contre la rage, marquant ainsi une avancée significative dans le domaine de la vaccination. Il est intéressant de noter que les vaccins peuvent agir à la fois de manière préventive et curative, notamment lorsque le temps d'incubation de la maladie est suffisamment long. Bien que dans de nombreuses parties du monde, les maladies infectieuses ne soient plus une menace majeure en raison d'une bonne couverture vaccinale, il est crucial de maintenir cette protection. La résurgence de maladies telles que la rougeole peut se produire lorsque la couverture vaccinale diminue, mettant en évidence la nécessité de maintenir un engagement constant en faveur de la vaccination pour prévenir ces maladies évitables.

En stimulant la réponse humorale, la vaccination favorise la production d'anticorps spécifiques, renforçant ainsi notre système immunitaire et nous protégeant contre les agents pathogènes. La réponse humorale est donc un élément essentiel de l'efficacité des vaccins.

3 - La réponse humorale

1 - La réaction inflammatoire

La réaction inflammatoire constitue la réponse dynamique et complexe du système immunitaire face à diverses agressions, qu'elles proviennent de l'extérieur, telles qu'une infection, une brûlure ou une blessure, ou de l'intérieur, comme la présence de cellules tumorales. Les manifestations classiques de l'inflammation se traduisent par la rougeur, la chaleur, la tuméfaction et la douleur. Lorsqu'un traumatisme ou une infection survient, des cellules clés du système immunitaire entrent en action, notamment les mastocytes et les macrophages, qui sont présents dans les tissus. Ces cellules jouent un rôle crucial en détectant les éléments associés à la présence de microorganismes (MAMP) ou liés à des dommages cellulaires (DAMP). Leur capacité de reconnaissance repose sur des récepteurs spécialisés appelés PRR (Pattern Recognition Receptors), qui identifient des profils moléculaires spécifiques. En réponse à cette reconnaissance, ces cellules immunitaires libèrent une variété de substances chimiques médiateurs de l'inflammation, parmi lesquelles les cytokines occupent une place centrale.

2 - L'immunité innée et adaptative

L'immunité, un système complexe et sophistiqué du corps humain, est divisée en deux composantes fondamentales : l'immunité innée et l'immunité acquise. Ces deux branches travaillent en tandem pour fournir une protection efficace contre les agents pathogènes, contribuant ainsi à maintenir la santé et à prévenir les maladies.

L'immunité innée constitue la première ligne de défense du corps contre les envahisseurs potentiels. Présente dès la naissance, elle offre une protection immédiate, mais peu spécifique. Cette forme d'immunité repose sur des barrières physiques, telles que la peau et les muqueuses, ainsi que sur des mécanismes cellulaires et moléculaires préexistants. La reconnaissance des motifs moléculaires communs aux pathogènes permet à l'immunité innée de réagir rapidement, mais de manière générale, face à une variété d'agents pathogènes.

À l'inverse, l'immunité acquise se développe au fil du temps en réponse à une exposition spécifique à des agents pathogènes. Cette forme d'immunité est caractérisée par sa spécificité et sa capacité à générer une mémoire immunologique. Lorsqu'une personne est exposée à un pathogène spécifique pour la première fois, le système immunitaire acquiert la capacité de le reconnaître et de monter une réponse plus rapide et plus ciblée en cas de réexposition ultérieure. Les cellules T et B spécialisées ainsi que la production d'anticorps sont des éléments clés de l'immunité acquise.

3 - Les cellules dendritiques

Les cellules dendritiques sont d'origine myéloïde. Le processus de maturation des cellules dendritiques les conduit généralement à migrer vers les tissus périphériques, où elles exercent une vigilance continue pour détecter d'éventuels antigènes. Ces cellules dendritiques, présentes dans les tissus, sont fréquemment désignées sous les termes de "cellules dendritiques immatures" ou "cellules dendritiques résidentes".

Les cellules dendritiques acquièrent leur maturité lors de l'activation de leurs récepteurs de reconnaissance de motifs moléculaires (PRR). Leur rôle prédominant consiste à capturer des microorganismes entiers ou sous forme de fragments, ainsi que d'autres antigènes, afin de les transporter vers les organes lymphoïdes secondaires. L'objectif de cette migration est d'amorcer la réponse adaptative. Les antigènes peuvent être présentés de deux manières distinctes. Dans le cas de la présentation aux lymphocytes B, les antigènes sont directement exposés à la surface des cellules dendritiques. Pour la présentation aux lymphocytes T, les antigènes sont internalisés par les cellules dendritiques et associés sous forme de peptides au complexe majeur d'histocompatibilité.

Les cellules dendritiques folliculaires (CDF) représentent une sous-population spécialisée de cellules dendritiques localisées dans les follicules lymphoïdes des ganglions lymphatiques, de la rate et d'autres organes lymphoïdes. Leur fonction principale est d'interagir étroitement avec les lymphocytes T folliculaires (Tfh) au sein des centres germinatifs, jouant ainsi un rôle essentiel dans la promotion de la réponse immunitaire adaptative.

4 - Les centres germinatifs

Les centres germinatifs (GCs), présents dans les organes lymphoïdes secondaires en réponse à une infection, se forment particulièrement au sein des centres des follicules des cellules B. Ces structures dynamiques émergent après la reconnaissance d'un antigène par des cellules B naïves, qui migrent ensuite vers la périphérie du follicule pour recevoir des signaux de co-stimulation de la part des Tfh. Cette interaction entraîne la prolifération des cellules B et la formation de clusters qui, en collaboration avec les cellules dendritiques folliculaires (FDCs), donnent naissance aux centres germinatifs primitifs.

Les centres germinatifs peuvent être divisés en deux zones distinctes sur le plan structurel et fonctionnel, connues sous le nom de zone sombre (DZ, de l'anglais dark zone) et zone claire (LZ, de l'anglais light zone). La DZ est le site où les lymphocytes B, désignés sous le nom de centroblastes, connaissent une prolifération massive tout en subissant l'hypermutation somatique (HMS), entraînant l'accumulation de mutations dans la région variable de leur récepteur des cellules B (BCR). La LZ abrite quant à elle les centrocytes, qui passent par une étape de sélection d'affinité de leur BCR pour un antigène, un processus appelé sélection clonale (Victoria et Nussenzweig, 2022). La sélection clonale favorise la survie des cellules dotées de BCR présentant la plus grande affinité pour l'antigène, tandis que celles dotées de BCR à affinité plus faible sont éliminées. C'est également dans la zone claire que les réactions de commutation de classe ont lieu.

5 - Les cellules T

La différenciation des lymphocytes T CD4⁺ en lymphocytes auxiliaires, également désignés sous le terme de lymphocytes Th (T helper), représente un élément central de la réponse immunitaire adaptative. Ces lymphocytes sont activés suite à la reconnaissance de l'antigène par leur récepteur TCR à la surface d'une cellule dendritique mature, ce qui les conduit à se différencier en diverses sous-populations, notamment les lymphocytes T auxiliaires (Th), les lymphocytes T mémoires, les Th folliculaires (Tfh) et les lymphocytes T régulateurs (Treg). Il est important de souligner le rôle spécifique des lymphocytes Tfh dans les centres germinatifs, où ils établissent un contact direct avec les lymphocytes B, dans la zone claire. En outre, leur sécrétion

de cytokines telles que l'IL-21 contribue de manière significative à la différenciation terminale des lymphocytes B en plasmocytes et en cellules mémoire.

6 - L'activation des lymphocytes B

Pendant leur maturation dans la moelle osseuse, les lymphocytes B développent des récepteurs de surface appelés BCR. Une fois formés, ils quittent la moelle osseuse et circulent dans le sang vers les organes lymphoïdes secondaires tels que les ganglions lymphatiques, la rate et les amygdales. On peut noter que la majorité des lymphocytes B résidant à l'intérieur des follicules des organes lymphoïdes secondaires sont des lymphocytes B folliculaires. Ces lymphocytes B reconnaissent des antigènes par leur récepteur BCR qualifiés de « thymo-dépendants », lesquels induisent une réponse humorale efficace en provoquant, outre l'activation des lymphocytes B, celle des lymphocytes T. Il faut noter que certaines molécules peuvent déclencher la différenciation de certains groupes spécifiques de lymphocytes B en l'absence de lymphocytes T. Ces molécules sont désignées sous le terme d'antigènes thymo-indépendants, à l'exemple des polysaccharides bactériens.

Au sein des organes lymphoïdes secondaires (figure 3), les lymphocytes Th identifient le complexe CMH II-peptide via leur récepteur TCR. Cette interaction facilite la fixation de la molécule CD40L présente sur le lymphocyte Th à son récepteur CD40 présent sur le lymphocyte B. De plus, les cytokines (IL-6, IL-21) exprimées par les lymphocytes Th se lient à leurs récepteurs spécifiques en surface du lymphocyte B. La convergence de ces différents signaux originaires de ces interactions induit l'activation du lymphocyte B, qui se multiplie, engendrant ainsi la formation d'un clone de lymphocytes B exprimant un BCR identique, phénomène que l'on qualifie d'expansion clonale.

Les lymphocytes B activés se différencient alors en lymphocytes B mémoire, qui réagiront de manière accélérée lors d'une prochaine rencontre avec l'antigène, en lymphocytes B migrant vers les organes lymphoïdes secondaires pour former des centres germinatifs, et en plasmocytes producteurs d'anticorps. Les plasmocytes, plasmablastes et plasmocytes à longue durée de vie représentent des cellules B spécialisées dans la synthèse d'anticorps. Les plasmablastes, produits par la prolifération rapide des cellules B activées, ont une durée de vie brève et jouent un rôle essentiel dans la réponse immunitaire immédiate en générant des anticorps de

manière intensive. Certains plasmablastes se spécialisent ensuite en plasmocytes à courte durée de vie, maintenant une production d'anticorps intense sur une période limitée. D'autres plasmablastes évoluent en plasmocytes à longue durée de vie, résidant principalement dans la moelle osseuse, et garantissent une production continue d'anticorps pour une immunité à long terme.

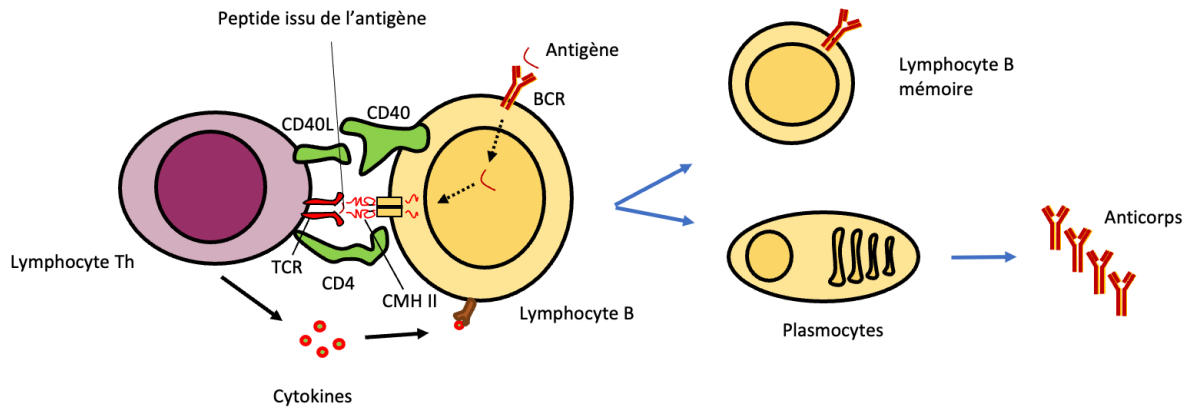


Figure 3: Activation des lymphocytes B.

La cellule dendritique (CDF) expose l'antigène natif sous forme de complexe immun au lymphocyte B (non représenté). L'antigène est reconnu par le BCR, puis internalisé et présenté sous forme de peptide par le CMH de classe II au lymphocyte T helper (Th). Cette interaction conduit à l'activation par des cytokines, entraînant la production de plasmocytes responsables de la sécrétion d'anticorps, ainsi que la formation de cellules B mémoire.

7 - Les immunoglobulines

La molécule emblématique de l'immunité adaptative est l'anticorps, ou immunoglobuline, qui possède une structure en forme de Y et est capable de se lier spécifiquement à des antigènes. L'histoire de sa découverte remonte à 1888, lorsque Jules Héricourt et Charlet Richet ont réalisé une observation cruciale. Leur travail a démontré que le sérum d'un animal immunisé pouvait conférer une partie de son immunité à un autre animal. Cette découverte a jeté les bases de la compréhension de l'immunité adaptative et a ouvert la voie à de futures avancées.

Au XIXe siècle, les deux principaux composants du sérum ont été identifiés comme l'albumine et les "globulines". En 1930, Tilenus a réalisé une observation significative lorsqu'il a noté qu'une fraction spécifique des globulines était enrichie après l'immunisation. Il a nommé cette fraction "fraction gamma", et ce terme a finalement donné son nom aux immunoglobulines IgG, qui sont les plus abondantes dans le sérum. Au fil du temps, la recherche a abouti à la découverte d'autres types

d'immunoglobulines. Les IgM ont été identifiées, suivies par les IgA, les IgD et les IgE. Chacune de ces classes d'immunoglobulines joue un rôle spécifique dans la réponse immunitaire adaptative.

Les IgG, abondantes dans le sérum, sont impliquées dans la protection à long terme contre les infections. Elles peuvent traverser le placenta, fournissant ainsi une immunité au fœtus.

Les IgA sont présentes principalement dans les sécrétions muqueuses et fournissent une protection contre les infections au niveau des muqueuses.

Les IgD ont un rôle moins compris, mais on les trouve sur la surface des lymphocytes B.

Les IgE sont associées aux réponses allergiques et à la défense contre les parasites. L'identification et la caractérisation de ces différentes classes d'immunoglobulines ont grandement enrichi notre compréhension de la façon dont le système immunitaire adaptatif fonctionne et réagit aux différentes menaces pour la santé.

4 - Les mécanismes de diversification des Immunoglobulines

Le récepteur de ces cellules B (Figure 4) est une immunoglobuline (Ig) membranaire composée de deux chaînes légères (IgL) et de deux chaînes lourdes (IgH) identiques. Chaque chaîne possède une région variable (V) et une région constante (C). Le site de reconnaissance de l'anticorps pour l'antigène est formé par la juxtaposition des régions variables V_H et V_L . La fonction effectrice des anticorps est par contre définie uniquement par la région constante de la chaîne lourde (C_H), qui définit l'isotype (IgM, IgG, IgE ou IgA).

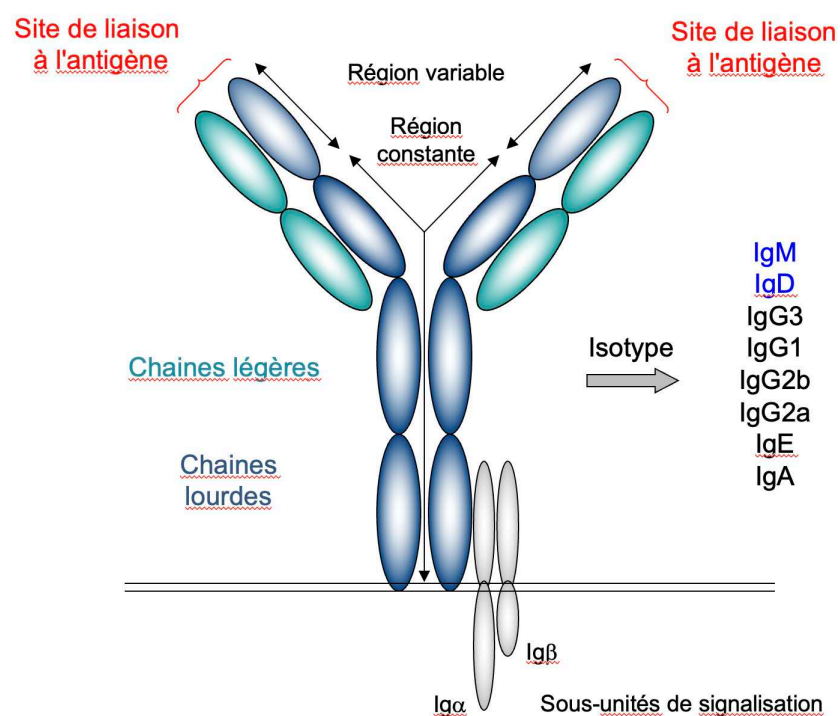


Figure 4: Structure du récepteur des cellules B.

Le récepteur des cellules B (BCR) est formé par l'association de deux chaînes lourdes et de deux chaînes légères, chacune hébergeant une région variable et une région constante. Les parties variables des chaînes lourdes et légères forment le site de liaison de l'antigène, tandis que les parties constantes des chaînes lourdes déterminent l'isotype et la fonction effectrice de l'anticorps. Le BCR est exprimé à la surface des cellules B et s'associe à l'hétérodimère $Ig\alpha/Ig\beta$, permettant la transduction du signal suite à la reconnaissance d'un antigène.

La diversité dans le répertoire des cellules B (Figure 5) est générée en deux grandes étapes (Rajewsky 1996). La première a lieu lors du développement des cellules B dans la moelle osseuse, où le mécanisme de recombinaison V(D)J assemble les régions variables V_H et V_L , de manière à ce que chaque cellule possède un récepteur à l'antigène (BCR) de spécificité unique. Cette recombinaison commence par la coupure

des segments génétiques V, D et J par les enzymes RAG-1 et RAG-2. Ces enzymes créent des cassures double brin à des endroits spécifiques du locus des récepteurs antigéniques. Ensuite, les segments V, D et J sont réarrangés de manière aléatoire et associés les uns aux autres pour former un gène complet de récepteur antigénique. Ces cellules qui possèdent un BCR de faible affinité et non spécifique des antigènes du soi, sortent en périphérie pour établir le répertoire primaire de cellules B. Ce répertoire permet à l'hôte de répondre à un grand nombre d'antigènes différents.

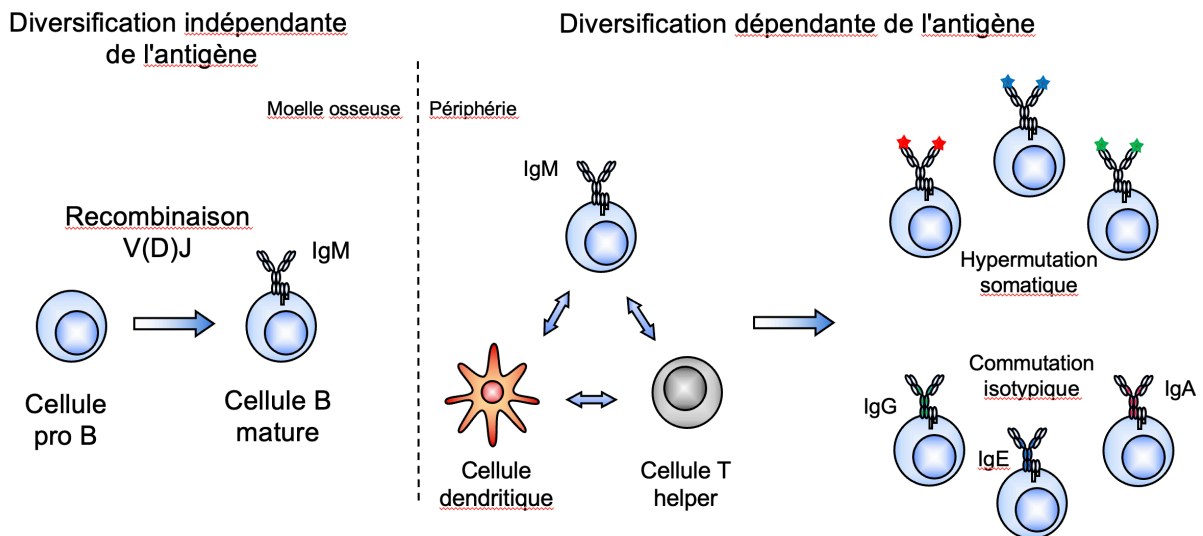


Figure 5: Mécanismes de génération de la diversité des anticorps.

Le répertoire primaire d'anticorps prend naissance dans la moelle osseuse par le biais de la recombinaison V(D)J, de manière indépendante de l'antigène, lors du développement des cellules B à partir de précurseurs hématopoïétiques (pré-pro B). Lors des réponses immunes dans les organes lymphoïdes secondaires, les cellules B activées par l'antigène enrichissent ce répertoire grâce aux mécanismes d'hypermutation somatique et de commutation isotypique.

Malgré sa grande diversité, ce répertoire est encore diversifié, dans une deuxième étape, au cours des réponses immunes par deux mécanismes : l'hypermutation somatique (HMS) et la commutation isotypique (CI). L'HMS (Figure 6) modifie l'affinité du récepteur pour l'antigène en induisant des mutations dans les régions V_H et V_L (Di Noia, 2007), alors que la CI (Figure 7) remplace l'isotype exprimé (de IgM à IgG, IgE ou IgA) par un événement de recombinaison ayant lieu au locus IgH et modifiant ainsi la fonction effectrice des anticorps, tout en préservant sa spécificité (Chaudhuri, 2007).

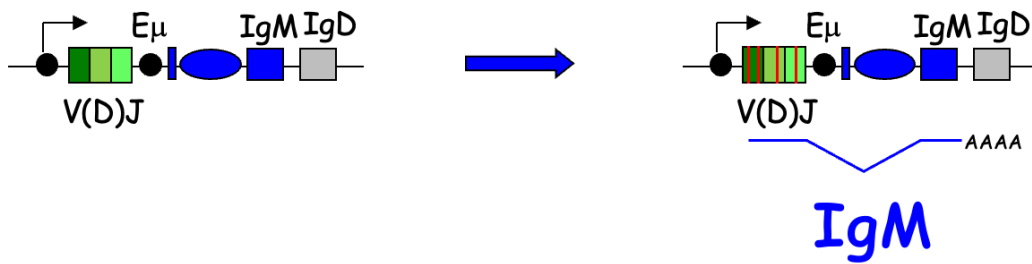


Figure 6: L'hypermutation somatique.

L'hypermutation somatique induit des altérations de la séquence d'ADN dans la partie variable réarrangée (VDJ) des chaînes IgH et IgL des anticorps. Ce mécanisme engendre des mutations ponctuelles, de petites insertions ou de petites délétions, toutes représentées par la zone rouge.

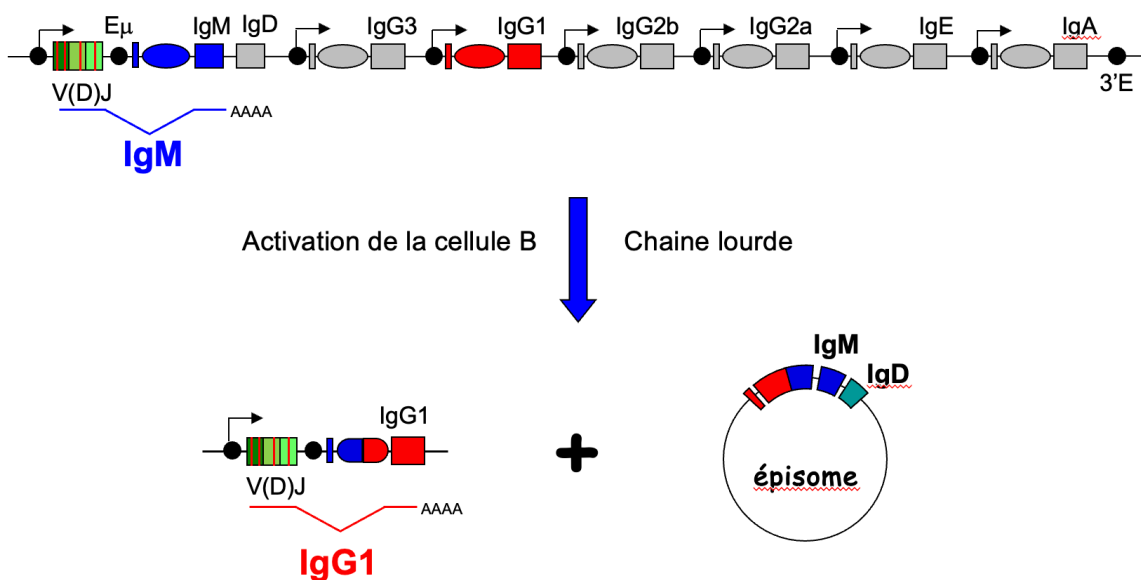


Figure 7: La commutation isotypique.

La commutation isotypique est une réaction de recombinaison qui se déroule au sein du locus IgH et vise à modifier l'isotype des anticorps. Cette recombinaison est spécifique à la région car elle survient entre deux zones dotées de séquences répétées appelées régions de switch (S ou ovale). Ces régions de switch font partie d'unités transcriptionnelles comprenant un promoteur induit par des cytokines (marqué d'un cercle noir) et un exon codant pour un nouvel isotype (C ou rectangle). Dans l'exemple présenté ici, la recombinaison a lieu entre la région donneuse S_μ et la région accepteuse S_{G1}. Les séquences entre ces deux régions de switch, impliquées dans la recombinaison, sont excisées du génome sous forme d'un épisode circulaire. Ainsi, l'exon C_μ est remplacé par C_{G1}, ce qui permet l'expression de l'isotype IgG1, sans toutefois altérer la région variable de l'anticorps.

La maturation d'affinité est un mécanisme biologique essentiel au sein du système immunitaire, principalement localisé dans les centres germinatifs des follicules lymphoïdes. Lorsqu'un lymphocyte B est exposé à un antigène, il génère initialement des anticorps avec une affinité de base. Toutefois, pendant la phase de prolifération

cellulaire, des mutations génétiques aléatoires surviennent, créant une diversité de BCR. Une sélection favorise les lymphocytes B qui produisent des anticorps dotés d'une affinité améliorée pour l'antigène. Ces cellules sélectionnées évoluent en plasmocytes, générant des anticorps de haute affinité. Ce processus d'affinement de la reconnaissance antigénique optimise la réponse immunitaire, renforçant progressivement la spécificité et l'efficacité des anticorps au fil du temps.

5 - **Activation-Induced cytidine Deaminase (AID)**

Les mécanismes de commutation isotypique et d'hypermutation somatique se rejoignent par leur dépendance à l'expression et à l'activité d'une même enzyme, appelée activation-induced cytidine deaminase (AID). La mise en évidence de l'implication d'AID dans l'initiation de ces processus, datant des années 2000, a constitué une percée majeure dans la compréhension moléculaire des mécanismes de diversification des anticorps liés aux antigènes.

1 - Découverte et fonction

La mise au jour d'AID a été facilitée par l'exploitation d'un modèle cellulaire dédié à l'investigation de la commutation isotypique (CI). Cette lignée cellulaire, identifiée sous le nom de CH12F3 (CH12), provient d'un lymphome B de souris et a été sélectionnée en raison de sa capacité à réagir de manière efficace au mélange de stimulations TGF- β , IL-4 et CD-40L, favorisant ainsi une commutation isotypique vers IgA (Nakamura, 1996).

Ainsi, un criblage par hybridation soustractive d'ADNc, réalisé à partir de CH12 stimulées et non-stimulées afin d'induire la commutation isotypique, a permis d'identifier AID comme étant exclusivement présente dans les CH12 stimulées.

De plus, il semble que son expression soit spécifiquement restreinte aux cellules B des centres germinatifs in vivo (Muramatsu, 1999). Les résultats de la génération ultérieure de souris déficientes pour AID ont mis en lumière son rôle essentiel à la fois dans la commutation isotypique (CI) et dans l'hypermutation somatique (HMS) (Muramatsu, 2000). De manière concomitante, l'équipe d'A. Durandy a établi un lien entre le syndrome hyper-IgM de type II et une déficience en AID (Revy, 2000).

Ces résultats mettent en évidence qu'AID constitue le facteur essentiel nécessaire pour les deux processus de diversification des immunoglobulines que sont la commutation isotypique (CI) et l'hypermutation somatique (HMS).

2 - Structure

AID est une petite protéine de 198 acides aminés (Figure 8) faisant partie de la famille des apolipoprotein B mRNA editing enzyme catalytic polypeptide-like (APOBECs). Elle

comporte un domaine catalytique situé entre les acides aminés 56 et 94, qui lui confère sa capacité de désaminer les cytosines. Par ailleurs, elle est dotée d'un signal de localisation nucléaire (NLS, pour nuclear localization signal) dans son domaine N-terminal (N-ter) et d'un signal d'export nucléaire (NES, pour nuclear export signal) dans son domaine C-terminal (C-ter).

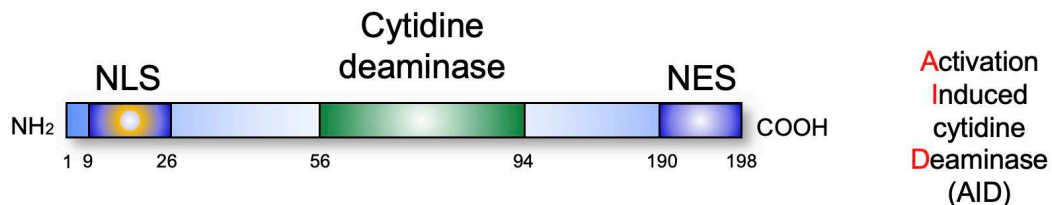


Figure 8: Activation induced cytidine deaminase (AID).

La séquence d'AID est conservée entre différentes espèces, et son domaine catalytique est caractérisé par un motif de déaminase de cytosine (entre les acides aminés 56 et 94). C'est ce domaine qui permet à l'AID de convertir spécifiquement des cytosines en uraciles dans l'ADN des gènes d'immunoglobulines. AID est pourvu d'un signal de localisation nucléaire (NLS) à son extrémité N-terminale, associé à un domaine essentiel pour l'hypermutation somatique (HMS). À son extrémité C-terminale, AID présente un signal d'export nucléaire (NES) ainsi qu'un domaine nécessaire à la commutation isotypique (CI). Les chiffres spécifiés correspondent aux positions des acides aminés délimitant chaque domaine.

3 - AID et le cancer

AID désamine des cytosines en uraciles dans les gènes des immunoglobulines (Figure 9), générant ainsi des mutations et des cassures double brin à l'ADN qui sont essentielles pour générer une très grande diversité d'anticorps.

Toutefois, malgré son rôle clé dans l'immunité humorale, AID possède un potentiel oncogénique important qui repose sur sa capacité à provoquer des dommages à l'ADN en dehors des loci des immunoglobulines. Ainsi, l'expression transgénique constitutive et ubiquitaire d'AID entraîne, entre autres, l'apparition de lymphomes T et de carcinomes (Okazaki, 2003). De plus, les lymphomes humains impliquant les lymphocytes B matures sont généralement associés à des translocations impliquant le locus IgH et des proto-oncogènes (Ramino, 2007). D'autre part, l'expression ectopique d'AID a également été impliquée dans la tumorigenèse, car l'expression d'AID peut être induite dans d'autres types cellulaires par des pathogènes tels que des virus oncogènes (Okazaki, 2007). En dépit des efforts, les mécanismes qui contrôlent la réparation efficace des lésions induites par AID, et qui préviennent les mutations généralisées et l'instabilité génomique, restent mal compris.

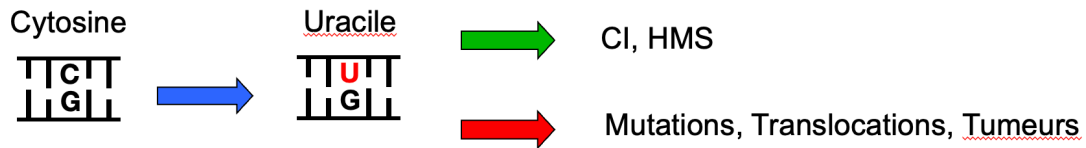


Figure 9: AID et cancer.

AID peut être surexprimée dans certains types de lymphomes, des cancers qui affectent les cellules du système immunitaire. Il peut également induire des mutations dans des gènes non liés aux immunoglobulines. Lorsque cette activité se produit dans des gènes suppresseurs de tumeurs ou des oncogènes, elle peut contribuer à l'initiation ou à la progression du cancer. AID peut également être exprimée dans certaines cellules non lymphoïdes et être impliquée dans le développement de cancers solides.

4 - Le modèle de déamination de l'ADN

Le groupe de M. Neuberger a élaboré un modèle explicatif novateur, présenté peu après la découverte d'AID, qui éclaire la manière dont les désaminations induites par AID dans l'ADN peuvent simultanément donner lieu à des mutations lors de l'hypermutation somatique (HMS) et à des cassures double-brin (DSB) dans le contexte de la commutation isotypique (CI). Ce modèle intègre des composants des voies de réparation par excision de base (BER, Base Excision Repair) et des éléments de la voie de réparation des mésappariements (MMR, mismatch repair) pour corriger les uraciles introduits dans l'ADN par AID (Figure 10). Ces voies de réparation, habituellement fidèles, sont intentionnellement détournées pour susciter délibérément des mutations ou des cassures double-brin lors du processus de diversification des immunoglobulines. L'étude approfondie du profil de mutation en lien avec la disparition de l'enzyme UNG (uracil-DNA glycosylase) a permis de démêler les mécanismes de mutagenèse et de réparation de l'ADN découlant des mutations induites par AID, contribuant ainsi à la compréhension des processus de recombinaison et de mutation nécessaires à la commutation isotypique et à l'hypermutation somatique (Petersen-Mahrt, 2002a ; Rada, 2002b).

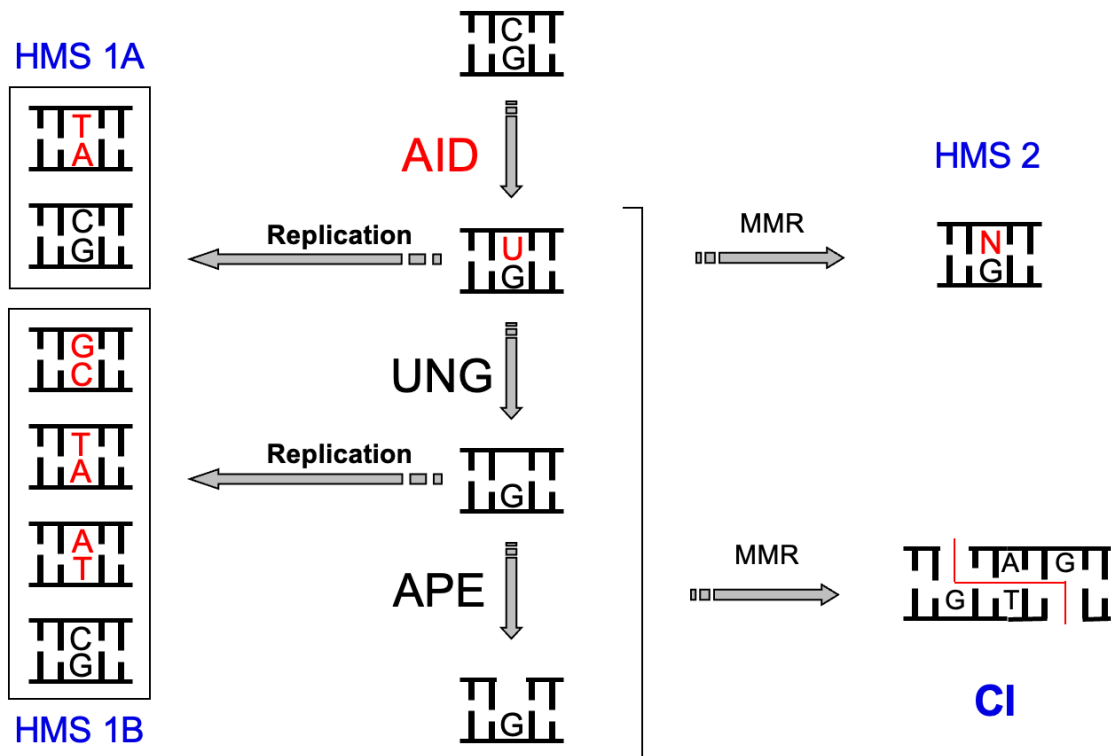


Figure 10: Le modèle de désamination de l'ADN.

AID induit la conversion des cytosines (C) en uraciles (U) au sein de l'ADN des régions V et S des loci des immunoglobulines. La réplication sans réparation préalable des U générés peut conduire à l'émergence de mutations, principalement de type transitions (HMS 1A). Les uraciles (U) générés dans l'ADN peuvent être identifiés par la voie de réparation des mésappariements (MMR), déclenchant ainsi la formation d'une cassure sur l'un des deux brins. Cette cassure est ensuite réparée par une polymérase non-fidèle, qui introduit principalement des transitions et des transversions aux bases A/T (HMS 2). L'uracile peut également être détecté et glycosylé par l'uracile-DNA glycosylase (UNG), créant ainsi un site abasique propice à la réplication et pouvant entraîner l'apparition de transitions et de transversions (HMS 1B). Ce site abasique peut également être traité par une endonucléase apurinique/apyrimidique (APE), provoquant ainsi la formation d'une cassure simple brin qui peut être réparée de manière précise par une ADN polymérase. La cassure simple brin peut finalement être convertie en cassure double brin (DSB) à extrémités sortantes si une autre cassure simple brin proche, située sur le brin opposé, est présente. Cette DSB est réparée par la jonction non homologue de l'extrémité libre (NHEJ, pour l'anglais "Non-Homologous End Joining") pour accomplir la recombinaison nécessaire à la commutation isotypique (CI).

5 - Les complexes protéiques régulateurs de AID

La régulation d'AID constitue un réseau sophistiqué d'interactions protéiques et moléculaires qui garantissent une régulation fine de son activité, permettant ainsi de maintenir un équilibre délicat entre la génération de diversité des anticorps et la préservation de l'intégrité génomique. Ces complexes régulateurs veillent à ce que l'activité d'AID soit précisément déployée dans les contextes appropriés du système immunitaire. AID établit des interactions avec des complexes nucléosomes, tels que les protéines histones, facilitant ainsi son accès aux régions cibles du génome. La

régulation de la localisation de AID est également cruciale, impliquant des protéines de transport telles que l'exportin 1 (XPO1). Des composants essentiels de ce réseau sont les protéines impliquées dans la réparation de l'ADN, notamment la protéine Uracil-DNA Glycosylase (Ung) et les protéines MutS Homologue 2 et 6 (Msh2 et Msh6), formant des complexes avec AID pour corriger les altérations induites par cette enzyme. Une autre dimension de cette régulation repose sur les protéines HP1, membres d'une famille de protéines intervenant dans la structure et la régulation de l'hétérochromatine, une forme plus compacte de la chromatine. Ces complexes sont d'une importance cruciale dans la maîtrise de l'activité de l'enzyme AID.

6 - Les cellules B et l'hypoxie

Le système vasculaire joue un rôle essentiel en fournissant de l'oxygène et des nutriments à tous les organes, tissus et cellules du corps, en fonction de leurs besoins métaboliques, tout en éliminant les déchets métaboliques tels que le dioxyde de carbone. Cela permet d'établir un équilibre entre l'approvisionnement et la consommation. La présence d'hypoxie dans un tissu est influencée par divers facteurs, tels que la taille et la densité des vaisseaux, la distance entre les cellules et le vaisseau le plus proche, la densité cellulaire, et les besoins métaboliques, créant ainsi des paysages hypoxiques distincts dans différents tissus de l'organisme (Figure 11). Dans des conditions physiologiques, la concentration en oxygène (en pourcentage) varie dans tout l'organisme, décroissant du sang artériel (~13 %) aux tissus lymphoïdes (~3 %, par exemple, la rate), en passant par les principaux organes (~6 %, par exemple, le cerveau, le foie, les poumons). Le processus de développement des progéniteurs hématopoïétiques en cellules B immatures se déroule dans la moelle osseuse (BM), où la présence de niches hypoxiques (moyenne de 2,25 %) est confirmée. Les différents niveaux d'hypoxie dans la moelle osseuse influent considérablement sur le destin et la fonction des cellules, y compris l'auto-renouvellement et le maintien de la quiescence dans les cellules souches, le recrutement de la lignée des cellules B, et le réarrangement V (D)J dans les cellules pro-B/pre-B (Chabi, 2019 ; Burrows, 2017). Les cellules B immatures passent ensuite dans une circulation bien oxygénée (moyenne de 13,2 %) avant d'atteindre divers tissus lymphoïdes secondaires. Là, elles rencontrent à nouveau des niches hypoxiques (rate 2,3 % et ganglions lymphatiques 1,85 %) et interagissent avec des antigènes spécifiques pour se différencier en plasmocytes et en cellules mémoire. Il est à noter que la majorité des centres germinatifs (CG) au sein des tissus lymphoïdes secondaires sont peu vascularisés, avec un niveau d'oxygène d'environ 1 %, et sont souvent situés à plus de 40 mm des vaisseaux sanguins. Ces centres germinatifs expriment fortement la protéine HIF (Facteur 1 Inductible par l'Hypoxie), essentielle à la formation de la diversité du répertoire d'anticorps et au développement d'une réponse immunitaire efficace (Abott, 2016). En outre, les cellules B activées subissent principalement une hypoxie dans la zone claire des centres germinatifs, bien que la valeur précise du niveau en oxygène

dans la zone claire/sombre des centres germinatifs n'ait pas été directement rapportée (Cho, 2016).

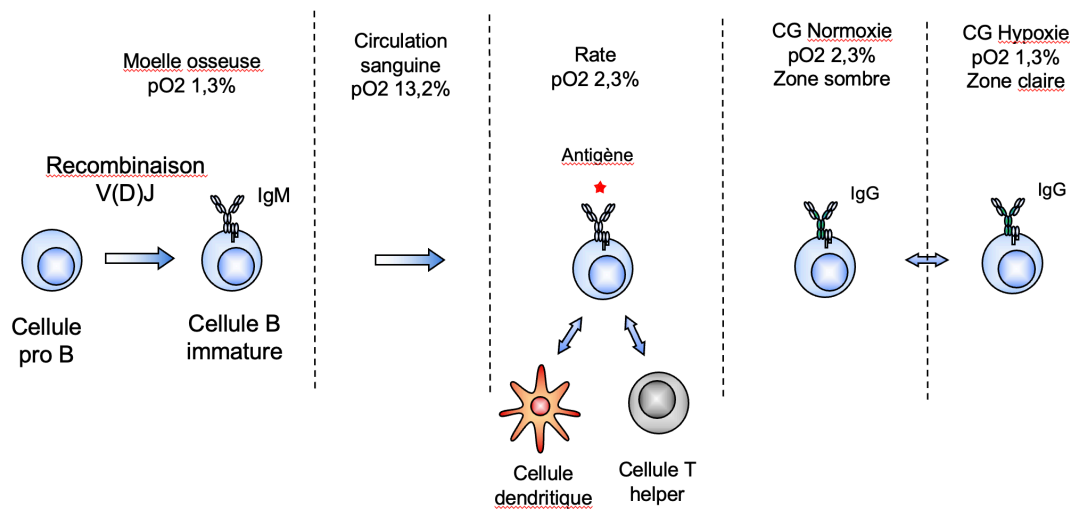


Figure 11: Tensions et gradients d'oxygène rencontrés au cours de la vie d'une cellule B.

Dans la moelle osseuse (MO), les cellules souches hématopoïétiques s'engagent dans la lignée des cellules B, donnant naissance à un grand nombre de cellules pro-B. Ces cellules pro-B subissent une recombinaison VDJ, aboutissant à la formation de cellules B immatures. La MO est caractérisée par un environnement hypoxique, avec une concentration en oxygène (pourcentage en oxygène : pO₂) d'environ 1,3 %, et présente des gradients hypoxiques extravasculaires variant de 0,6 à 2,8 %. Les cellules B immatures quittent la MO et entrent dans la circulation, où elles se présentent en tant que cellules transitionnelles, exposées à un environnement bien oxygéné avec une pO₂ de 13,2 %. Par la suite, elles migrent vers la rate pour finaliser leur maturation. Bien que la rate soit moins hypoxique que la MO, avec une pO₂ d'environ 2,3 %, il est probable qu'il existe des gradients d'oxygène à l'intérieur de la rate, variant de 0,5 à 4,5 %. Lorsqu'elles rencontrent un antigène, les cellules B subissent une activation et se différencient. Certains de ces lymphocytes B activés prolifèrent sous l'influence des cellules T et des cellules dendritiques folliculaires, formant un centre germinatif (CG). La plupart des CG se forment à une distance significative du vaisseau splénique, avec une pO₂ d'environ 1,3 %, suggérant que les gradients hypoxiques déterminent l'emplacement des CG. Ce mécanisme contrôle ainsi la colocalisation anatomique des cellules immunitaires et des antigènes, un aspect essentiel à l'établissement d'une réponse immunitaire efficace.

7 - Le complexe HIF

Le Facteur 1 Inductible par l'Hypoxie (HIF-1), découverte en 1992 par Semenza (Semenza, 2012) est impliquée dans un mécanisme sophistiqué par lequel les cellules réagissent et s'adaptent à des niveaux réduits d'oxygène dans leur environnement, un état appelé hypoxie. Cette protéine agit comme un chef d'orchestre moléculaire, coordonnant une réponse cellulaire précise pour faire face à des conditions d'oxygénation défavorables. Lorsque les cellules détectent une diminution des niveaux d'oxygène, le HIF-1 entre en action. Il fonctionne comme un facteur de transcription, ce qui signifie qu'il influence directement l'activité génique en régulant l'expression de certains gènes. Cette capacité de modulation génique permet aux cellules de s'adapter de manière rapide et spécifique aux changements dans leur environnement. Les fonctions primaires du HIF-1 sont multiples et cruciales. Tout d'abord, il favorise l'angiogenèse (VEGF, Vascular endothelial Growth factor), un processus qui implique la création de nouveaux vaisseaux sanguins. Cela est particulièrement important dans des conditions d'hypoxie, car cela augmente l'apport d'oxygène aux tissus, contribuant ainsi à maintenir la viabilité cellulaire. De plus, le HIF-1 régule la glycolyse anaérobie, un processus métabolique alternatif qui produit de l'énergie en l'absence d'oxygène. En favorisant la glycolyse, le HIF-1 permet aux cellules de maintenir une production d'énergie adéquate même dans des environnements faiblement oxygénés. Enfin, le HIF-1 coordonne d'autres adaptations cellulaires visant à améliorer la survie en conditions d'hypoxie. Ces adaptations peuvent inclure des changements dans le métabolisme cellulaire, la synthèse de protéines spécifiques et d'autres ajustements moléculaires.

1 - Structure des HIFs

HIF-1 est un hétérodimère composé de deux sous-unités distinctes, HIF-1 α (826 acides aminés) et HIF-1 β (789 acides aminés). Ces sous-unités présentent une organisation structurale complexe, comprenant trois domaines majeurs (Figure 12). Tout d'abord, un domaine prédominant constitué principalement de motifs bHLH (basic-helix loop helix) et PAS (Per-Arnt-Sim) est impliqué dans la dimérisation des protéines et dans la liaison du complexe à l'ADN. Ce domaine joue un rôle crucial dans la formation de la structure tridimensionnelle du complexe HIF-1 et sa capacité à

interagir avec les séquences d'ADN spécifiques. En partie C-terminale, plusieurs domaines de transactivation sont localisés, contribuant à la régulation de l'expression génique. Ces domaines facilitent l'activation transcriptionnelle du complexe HIF-1, régulant ainsi les réponses cellulaires à l'hypoxie. Enfin, la région ODD (oxygen-dependent degradation domain) est responsable de la dégradation d'HIF-1 α en conditions normales d'oxygène (normoxie) par le protéasome. Ce mécanisme de dégradation assure une régulation serrée de la concentration d'HIF-1 α dans la cellule, permettant une réponse adaptative aux variations de l'oxygène environnant. La sous-unité HIF-1 β possède également un domaine bHLH-PAS, essentiel pour former le complexe hétérodimère avec HIF-1 α . Il est important de noter l'existence de plusieurs isoformes de HIF-1 α , parmi lesquelles HIF-1 α et HIF-2 α ont été particulièrement étudiées en raison de leur rôle central dans la régulation des réponses cellulaires à l'hypoxie.

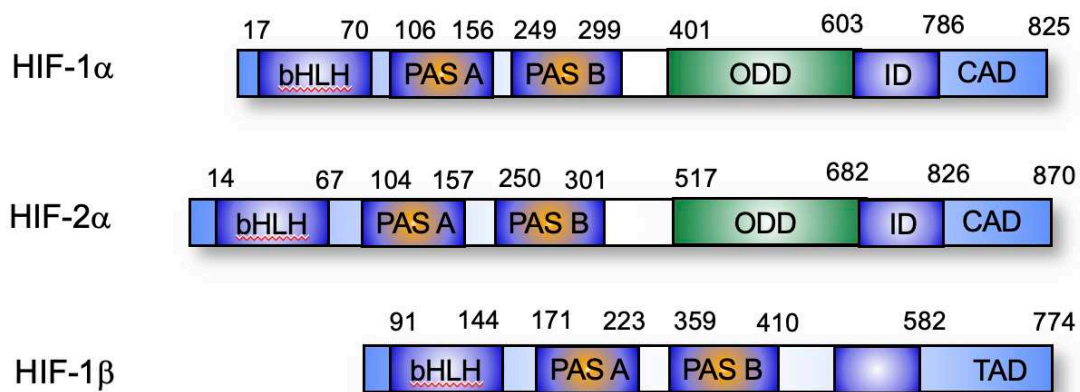


Figure 12: Structure des facteurs HIFs.

HIF-1 α , HIF-2 α et HIF-1 β font partie de la famille des protéines caractérisées par des motifs bHLH (basic helix-loop-helix), comprenant un domaine bHLH à l'extrémité N-terminale et deux domaines PAS (Per-ARNT-Sim). HIF-1 α et β présentent un domaine de dégradation sensible à l'oxygène (ODD), régulant ainsi leur stabilité, et un domaine de transactivation à l'extrémité C-terminale (CAD), qui est inhibé en conditions normoxiques par le domaine d'inhibition (ID). HIF-1 β dispose d'un domaine de transactivation (TAD), bien que ce dernier n'ait pas de fonction dans le contexte de l'activité des facteurs HIFs.

2 - Régulation des HIFs

Au cours de la réponse immunitaire adaptative, les lymphocytes B vont produire des immunoglobulines d'affinité croissante à l'antigène. Ce phénomène de maturation prend place dans des organes lymphoïdes secondaires au sein des centres germinatifs (CG). Or, ces centres sont connus pour être localisés dans un microenvironnement hypoxique, car très faiblement vascularisé (Abbott, 2016 ; Zhang,

2022). Qui dit hypoxie dit régulation par le complexe protéique HIF. On ne connaît que peu les mécanismes physiologiques de régulation de ces processus dans le microenvironnement des GC et il est possible que la tension d'oxygène et la réponse physiologique à l'hypoxie jouent un rôle clé dans la régulation de la réponse humorale au sein des GC (Cho, 2016 ; Burrows, 2017).

La réponse cellulaire à l'hypoxie est médiée par le complexe HIF, formé par les sous-unités HIF-1 α et HIF-1 β (Figure 13). En normoxie, la protéine HIF-1 α est instable, elle fixe le facteur suppresseur de tumeur von Hippel-Lindau (pVHL), une ubiquitine ligase E3 qui l'ubiquitine et déclenche sa dégradation par le protéasome. L'hypoxie enclenche la stabilisation de HIF-1 α , sa translocation nucléaire, son association avec HIF-1 β pour former le complexe HIF-1. Puis ce dernier s'associe à d'autres cofacteurs tels que p300/CBP pour se lier aux éléments de réponse à l'hypoxie (HRE) et induire la transactivation des gènes sensibles à l'hypoxie (Zhang, 2022). Le gène le plus connu est l'EPO (Érythropoïétine), une hormone glycoprotéine acheminée par voie sanguine vers la moelle osseuse pour augmenter le nombre de globules rouges dans l'organisme. C'est d'ailleurs l'étude du promoteur de l'EPO en situation d'hypoxie qui a permis de mettre en évidence la séquence du HRE. VEGF est un autre gène cible spécifique à l'hypoxie bien connu.

En outre, il est décrit depuis peu que la voie de signalisation HIF peut également être activée de manière indépendante de l'oxygène par l'activation du récepteur des cellules B, des récepteurs Toll-like ou des récepteurs de cytokines, ce qui entraîne la stabilisation de la protéine HIF-1 α dans les cellules B. Il a été également montré que HIF-1 α est capable de se lier au HRE dans le promoteur d'AID et d'induire la production de l'ARNm d'AID (Liu, 2013).

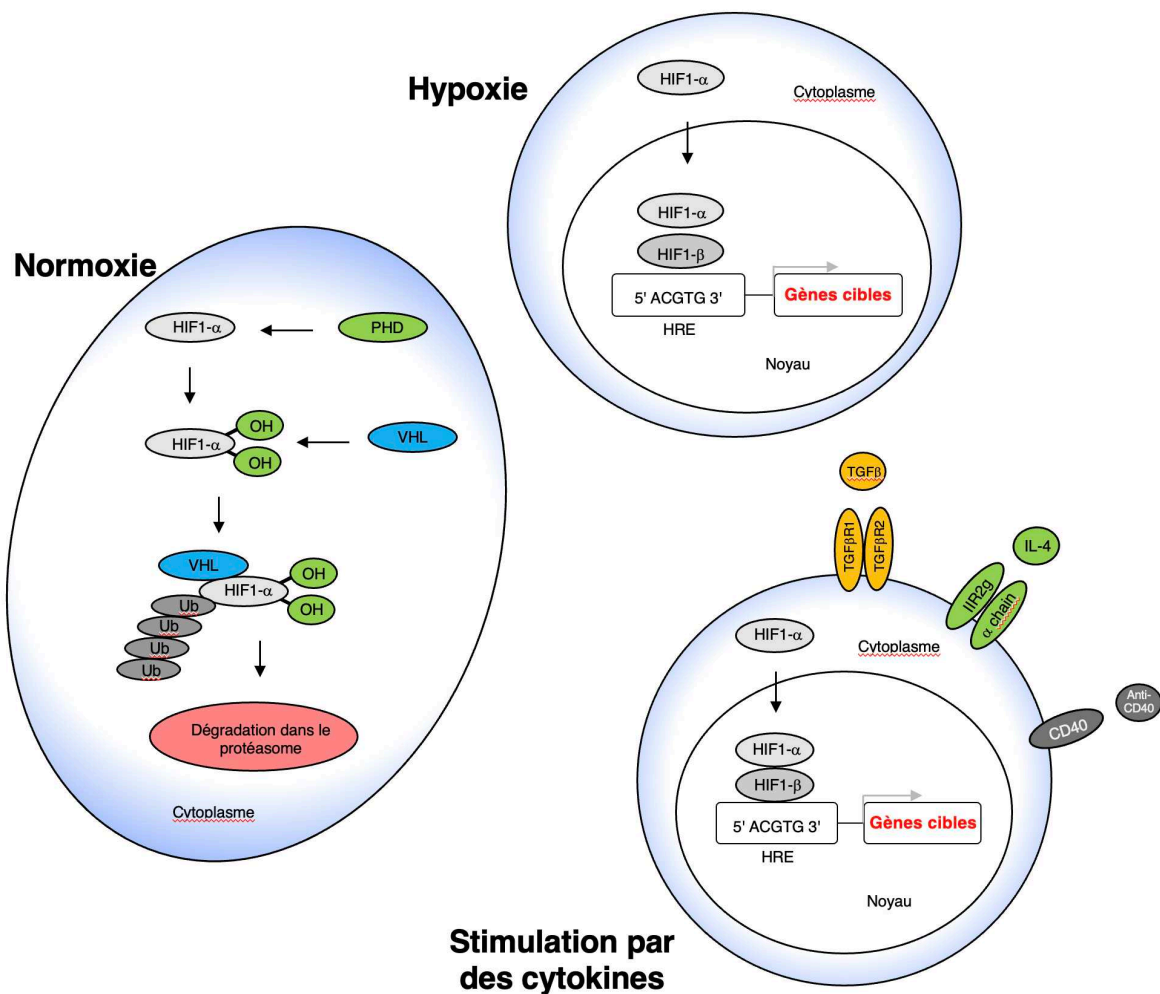


Figure 13: Mécanismes de détection de l'oxygène par les cellules et stimulation par les cytokines dans les cellules B (CH12).

Dans des conditions normoxiques, la protéine HIF-1 α , subit une hydroxylation sur deux résidus proline, par le prolyl hydroxylase (PHD). En conséquence, elle se lie à la protéine VHL, subit une ubiquitination, qui conduit à sa dégradation par le protéasome. En revanche, en conditions hypoxiques, ou sous stimulation de cytokines, HIF-1 α se transloque vers le noyau cellulaire où il forme un complexe avec HIF-1 β . Le complexe résultant se lie ensuite à la région HRE (Élément de Réponse à l'Hypoxie) de l'ADN, déclenchant ainsi la transcription de gènes impliqués dans divers processus biologiques.

Mon hypothèse est que HIF1 pourrait être impliqué dans la régulation de l'activation transcriptionnelle du gène AID dans les cellules B et donc jouer un rôle actif dans la commutation isotypique (CI) au cours des réponses immunitaires.

8 - Objectif de ma thèse

Les anticorps, jouant un rôle crucial dans la défense contre les infections et les maladies, sont synthétisés par les lymphocytes B. Nous avons observé que la diversité dans le répertoire des cellules B émerge à travers deux étapes majeures. La première a lieu lors du développement des cellules B dans la moelle osseuse, où le mécanisme de recombinaison V(D)J assemble les régions variables V_H et V_L . Les cellules qui possèdent un anticorps fonctionnel et de faible affinité pour les antigènes du soi, sortent en périphérie pour établir le répertoire primaire de cellules B. Ce répertoire permet à l'hôte de répondre à un grand nombre d'antigènes différents.

Malgré sa grande diversité, ce répertoire est encore diversifié, dans une deuxième étape, au cours des réponses immunes par deux mécanismes : l'hypermutation somatique (HMS) et la commutation isotypique (CI). Ensemble, ces deux mécanismes aboutissent à la génération de réponses humorales de haute affinité, adaptées à l'antigène et de longue durée.

Ces deux mécanismes dépendent de l'action de l'enzyme AID (activation-induced cytidine deaminase), qui désamine des cytosines en uraciles dans les gènes des immunoglobulines, générant ainsi des mutations et des cassures double brin à l'ADN qui sont essentielles pour générer une très grande diversité d'anticorps.

Toutefois, malgré son rôle clé dans l'immunité humorale, AID possède un potentiel oncogénique important qui repose sur sa capacité à provoquer des dommages à l'ADN en dehors des loci des immunoglobulines. En dépit des efforts, les mécanismes qui contrôlent la réparation efficace des lésions induites par AID, et qui préviennent les mutations généralisées et l'instabilité génomique, restent mal compris.

Compte tenu du fort potentiel oncogène de AID, son expression et son activité doivent être étroitement régulées. Au niveau transcriptionnel, l'expression de AID n'est induite que dans les cellules B activées et résulte d'un équilibre étroit entre l'activité de certains facteurs de transcription.

Identifier les protéines régulatrices d'AID permet de mieux comprendre comment l'équilibre entre sa fonction physiologique et pathologique est régulé. Ces protéines de régulation interviennent pour s'assurer que l'activité de AID est limitée aux lymphocytes B activés et aux régions spécifiques des gènes codant pour les anticorps,

en évitant ainsi des altérations génétiques inappropriées pouvant déclencher des cancers et immunodéficiences.

L'idée est d'identifier les protéines associées à AID par une approche protéomique, en les isolant par des réactions de co-immunoprécipitation, et en les identifiant par spectrométrie de masse. Ce sera la première partie de mes résultats.

Ce projet a conduit à l'identification de plus de 300 protéines, et, à la découverte et la caractérisation fonctionnelle de 4 complexes protéiques d'importance, impliqués dans la régulation d'AID.

L'étude fonctionnelle de ces protéines nécessitant la création de nombreux vecteurs d'expression, et afin de les créer rapidement, j'ai mis au point un système de clonage innovant (UES; Universal Expression System), que je développerai en deuxième partie.

Enfin, dans une troisième et dernière partie, je vais présenter mon sujet principal de recherche, qui est centré sur le rôle du complexe protéique HIF1 (Hypoxia Induced Factor 1). En effet, il a été démontré que durant les réponses immunitaires, des régions d'hypoxie sont observées au sein des GC et que la réponse à l'hypoxie module la réponse humorale. Néanmoins, on connaît peu les mécanismes physiologiques de régulation de ces processus dans le microenvironnement des GC et il est possible que la tension d'oxygène et la réponse physiologique à l'hypoxie jouent un rôle clé dans la régulation de la réponse humorale au sein des GC. J'ai ainsi montré que l'expression de AID et l'efficacité de la recombinaison de classe des immunoglobulines sont conditionnées par le facteur inductible par l'hypoxie.

Chapitre 2 - Résultats

1 - Identification des co-facteurs d'AID

Identifier les protéines régulatrices d'AID permet de mieux comprendre comment l'équilibre entre sa fonction physiologique et pathologique est régulée. Ces protéines de régulation interviennent pour s'assurer que l'activité de AID est limitée aux lymphocytes B activés et aux régions spécifiques des gènes codant pour les anticorps, en évitant ainsi des altérations génétiques inappropriées pouvant déclencher des cancers et immunodéficiences. L'idée est d'identifier les protéines associées à AID, en les isolant par une réaction de co-immunoprécipitation, pour par la suite conduire des expériences fonctionnelles.

Mon projet d'identification des protéines par co-immunoprécipitation et spectrométrie de masse requière la production à grande échelle de lignées cellulaires d'intérêt et la disponibilité d'anticorps spécifiques (Figure 14). Pour identifier les protéines associées à AID, j'ai utilisé la lignée cellulaire CH12. Cette lignée est un lymphome B murin qui est capable de réaliser la Commutation Isotypique CI (de IgM à IgA), de manière dépendante d'AID et à des fréquences très élevées (Nakamura, 1996).

Pour pouvoir immuno-précipiter AID et identifier les protéines associées, j'ai transduit les cellules CH12 à l'aide d'un rétrovirus coexprimant la protéine AID portant les épitopes Flag et HA (2T-AID), un rapporteur fluorescent (GFP; Green Fluorescent Protein), et un marqueur de sélection (résistance à la puromycine). J'ai sélectionné les cellules transduites (exprimant AID) en présence de puromycine, puis triées par cytométrie de flux en fonction de l'expression de GFP (FACS Aria BD). J'ai vérifié l'expression d'AID par Western blot à l'aide des anticorps spécifiques des épitopes Flag et HA.

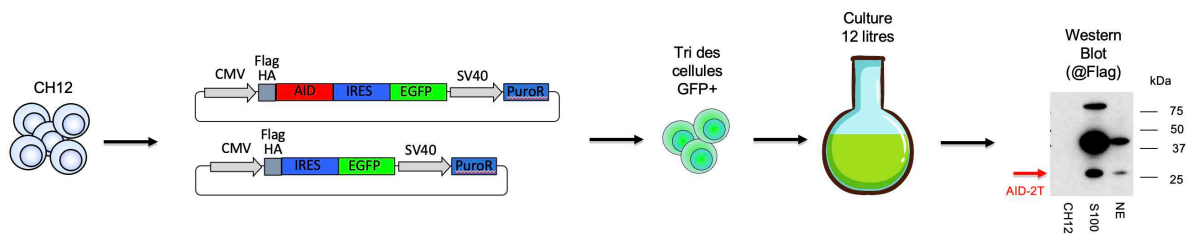


Figure 14: Création de la lignée CH12 sur exprimant AID.

Les cellules CH12 sont transduites avec un rétrovirus coexprimant la protéine AID-Flag-HA (ou non), le rapporteur GFP et la sélection à la puromycine. Les cellules GFP⁺ sont triées par FACS, puis cultivées en ballon de 12 litres. Les extractions cytoplasmiques et nucléaires sont testées par Western Blot (@Flag M2).

1 - Extraction des protéines cytoplasmiques et nucléaires

Le culot de cellules CH12 est repris dans le tampon de lyse S100 (voir Matériels et Méthodes). Après broyage des tissus (Wheaton douce tissue grinder, B) et centrifugation, le surnageant contenant les protéines solubles est récupéré (S100 ou extrait cytoplasmique). Le culot obtenu est lavé, puis repris dans le tampon de lyse NE. Après broyage du culot, incubation 1 heure sous agitation et centrifugation, on récupère le surnageant contenant les protéines nucléaires (NE, ou extrait nucléaire). S100 et NE sont dialysés contre le tampon BC100.

2 - Immuno-précipitation en tandem

L'immuno-précipitation (IP) est réalisé avec 60 mg d'extrait cellulaire (Figure 15). Il est incubé pendant quelques heures avec des billes d'agarose couplées avec l'anticorps anti-flag M2 (Sigma). Après lavages, les protéines immuno-précipitées sont éluées avec le peptide Flag (0,2ug/ul). Puis l'éluat est incubé quelques heures avec des billes agarose couplées avec l'anticorps anti-HA (Sigma). Après lavages, les protéines immuno-précipitées sont éluées avec un tampon glycine (Glycine 0,25 mM ; pH 2,5), puis neutralisées à pH7.

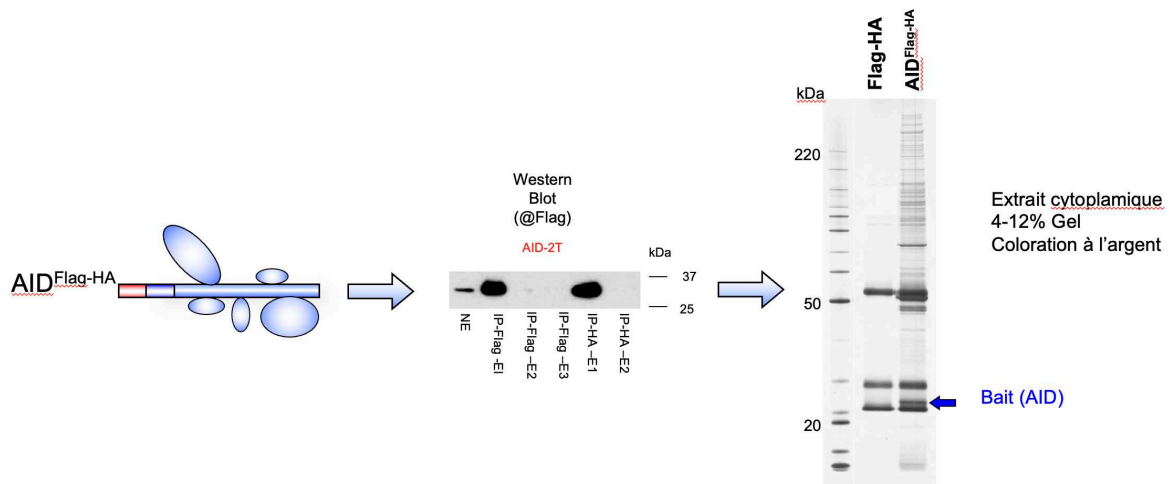


Figure 15: Identification des partenaires d'AID.

Une protéine de fusion a été créée : l'enzyme AID est couplée à une étiquette « Flag-HA ». Celle-ci est surexprimée dans des lymphocytes B de souris. Des extraits nucléaires sont préparés et AID est purifiée par immunoprécipitation grâce au tag Flag et HA. Les protéines sont séparées par électrophorèse sur gel (SDS-PAGE). La protéine AID est bien visible (Bait).

3 - Découverte du complexe KAP1, HP1 et AID

Pour purifier les complexes protéiques associés à AID j'ai conduit des co-immunoprécipitations consécutives à l'aide d'anticorps spécifiques des épitopes Flag puis HA. J'ai élué les complexes protéiques en conditions natives par compétition avec une solution contenant les peptides Flag ou HA. Après validation de l'immunoprécipitation d'AID par Western blot, j'ai séparé les complexes protéiques purifiés par électrophorèse (SDS-PAGE). J'ai découpé la piste du gel en 20 bandes, qui ont ensuite été digérées à la trypsine. Les peptides ainsi obtenus ont été soumis à une analyse par spectrométrie de masse (LC-MS/MS). C'est la méthode classique dite «Gel Based Analysis ». J'ai ainsi identifié une première trentaine de protéines. AID a été identifiée dans la bande correspondante à son poids moléculaire, validant ainsi l'analyse. J'ai montré la relevance de ces expériences par l'identification de protéines partenaires ayant un rôle fonctionnel dans le mécanisme de la commutation isotypique ou d'hypermutation somatique (Msh2, Msh6, XPO1) ainsi que d'autres n'ayant pas été caractérisées, comme par exemple la protéine KAP1. Cette protéine a retenu mon attention car elle fait partie d'un complexe (contenant les protéines HP1) impliqué dans la régulation de l'expression génique et car elle a été décrite comme impliquée dans la réparation de l'ADN.

J'ai validé l'identification de KAP1 par spectrométrie de masse par des expériences de co-immunoprécipitation réciproques suivies d'une analyse par Western blot. J'ai pu ainsi, montrer qu'AID est associé dans un même complexe à KAP1, dans les fractions nucléaires et chromatiniennes.

Pour mieux caractériser ce complexe, j'ai mis au point une méthode de purification adaptée à nos complexes protéiques en utilisant une colonne de filtration sur gel de Sépharose (exclusion de taille), en partenariat avec une équipe de Biologie Structurale (Figure 16). Après une immunoprécipitation du complexe, et j'ai chargé l'éluat sur une colonne de filtration (qui sépare les protéines en fonction de leur taille en conditions non-dénaturantes). Une analyse par Western Blot des différentes fractions m'a permis alors de démontrer que KAP1, HP1 et AID sont éluées dans les mêmes fractions. Par ailleurs, j'ai montré par cette caractérisation biochimique qu'AID fait partie d'un complexe d'un poids moléculaire supérieur à 670 kDa.

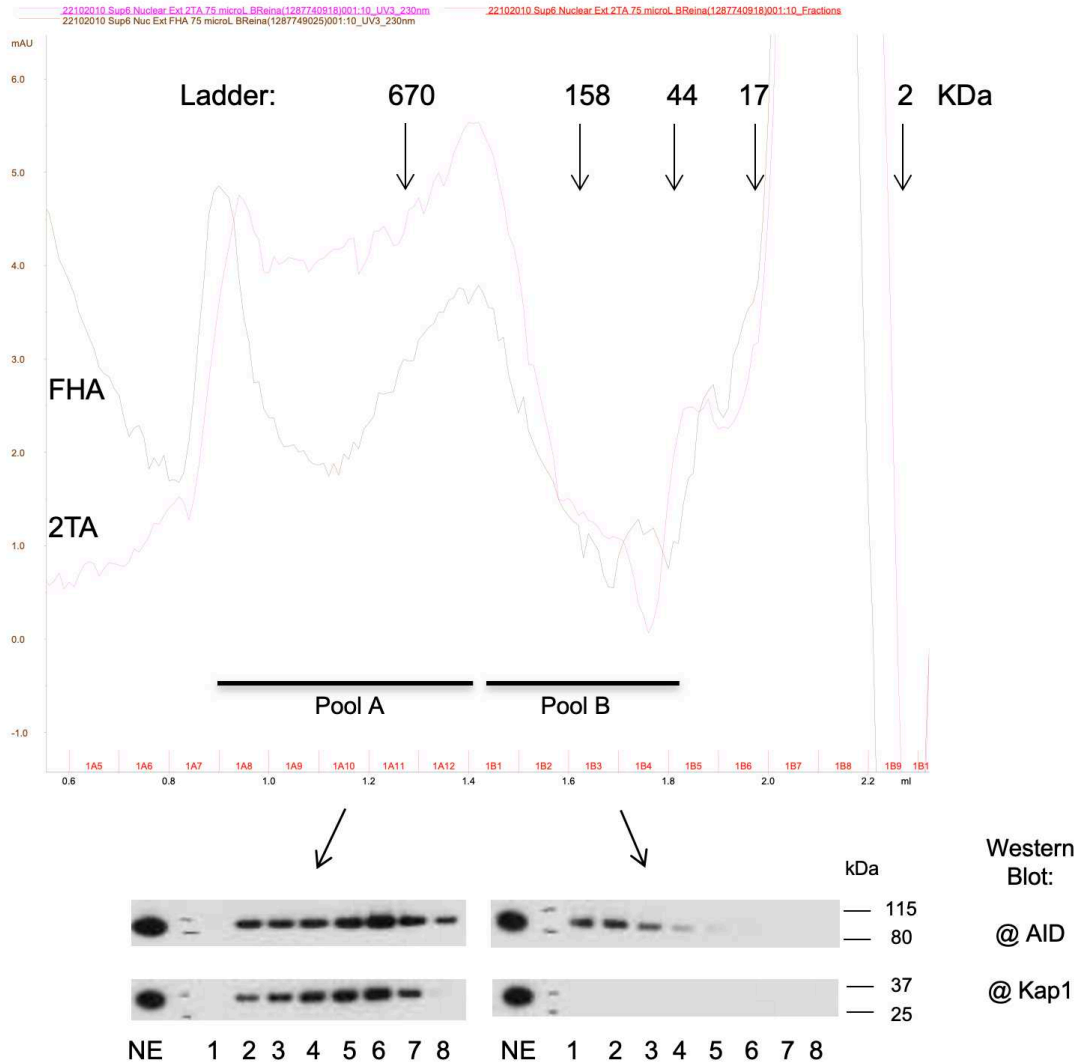


Figure 16: Chromatographie sur colonne de sépharose.

IP-Flag sur 20 mg, fraction NE, élution en tampon BC100-Chaps-peptide Flag P70 à 0,2 ug/ul. Gel filtration with AKTA micro Superose 6 PC3.2/30, colonne conditionnée avec le tampon BC100 chaps : 5000 to 5.10⁶ Da. 8 fractions ont été analysées, puis regroupées en un pool (A), et 8 autres fractions en un autre pool (B) pour analyse en spectrométrie de masse. Une majorité de protéine AID est détecté par Western Blot dans le pool A, ainsi que la protéine Kap1.

Par la suite, à l'aide des systèmes cellulaires et un modèle de souris knockout, nous avons montré que KAP1 est impliqué dans le recrutement d'AID aux locus IgH (Jeevan-Raj, 2011).

4 - Caractérisation des complexes protéiques associés à AID par MudPIT

Le haut poids moléculaire de ce complexe a montré qu'il est composé de diverses protéines, qui pourraient être importantes pour le contrôle de la fonction d'AID. J'ai

alors entrepris leur identification. J'ai réalisé une nouvelle préparation du complexe protéique (Figure 17). Après filtration sur gel, les fractions contenant AID et ses partenaires ont alors été soumis à la nouvelle technologie d'identification de protéines MudPIT (de l'anglais Multidimensionnel Protein Identification Technology), mise au point par la plateforme de spectrométrie de masse de l'IGBMC. Cette technologie, complémentaire de la LC-MS-MS, a également l'avantage d'être quantitative, car c'est une méthode dite de « shotgun protéomique » où l'on injecte l'ensemble de la digestion protéolytique, sans biais dû à une découpe de gel. Les purifications par co-immunoprécipitation et filtration sur gel, couplée à l'analyse par MudPIT, m'ont permis de multiplier par un facteur 10 la sensibilité de détection.

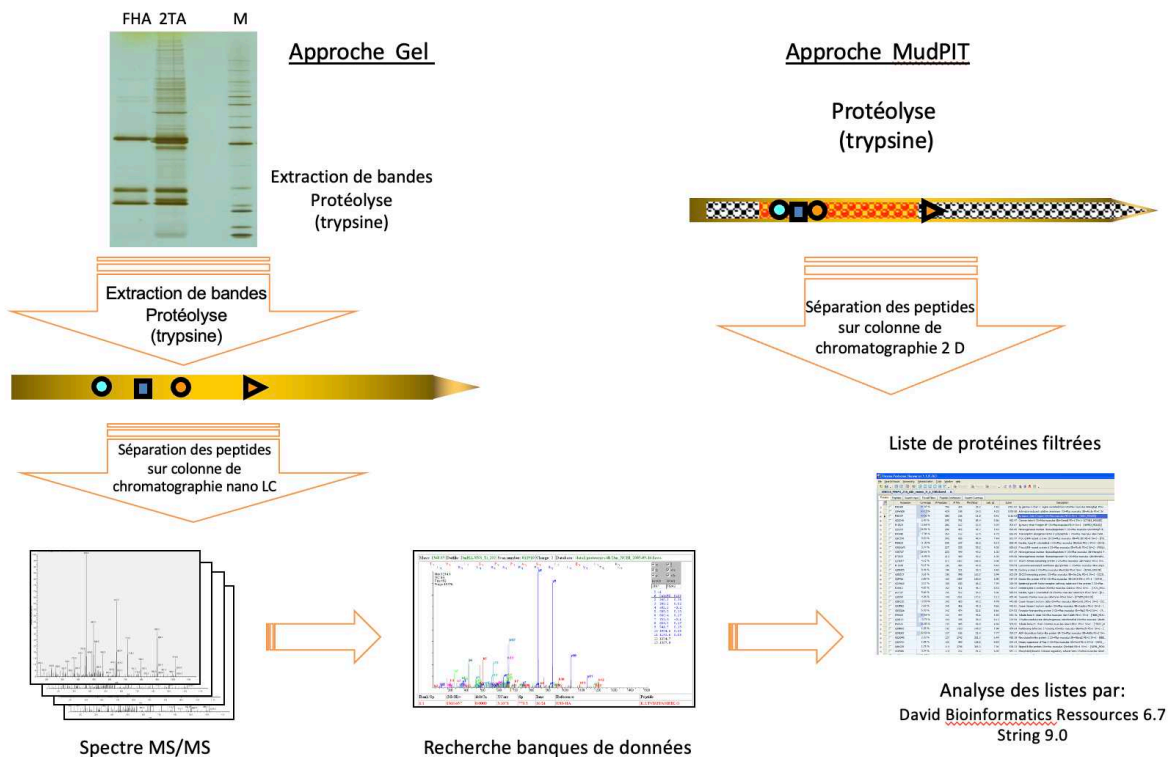


Figure 17: Approches GBA et MudPIT.

En approche gel, les bandes sont découpées, traitées par protéolyse. Les peptides sont séparés sur une colonne de résine C18 (séparation par degré d'hydrophobie croissant), couplé au spectromètre de masse (Orbitrap). Dans le cas de l'approche Mudpit, les peptides sont séparés sur une colonne 2D. La première partie est une résine SCX qui permet de séparer les peptides par un gradient de sels, puis une partie en résine C18.

Au final, plus de 300 protéines potentiellement impliquées dans la régulation fonctionnelle d'AID ont été identifiées (Figure 18). L'utilisation des logiciels informatiques Manteia, Sting 9 et Uniprot m'a permis d'identifier *in silico* plusieurs

sous-complexes ayant potentiellement un lien avec AID dont les complexes PAF, Cohésine et Mediator, que je vais brièvement détailler maintenant.

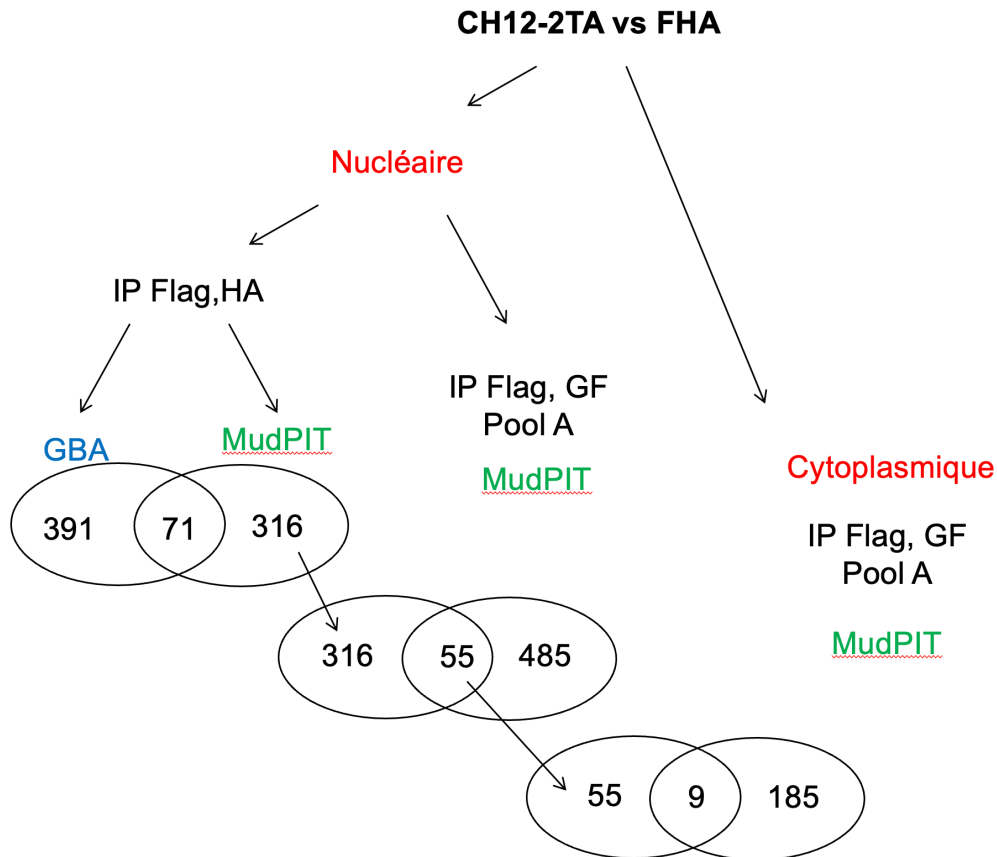


Figure 18: Bilan des protéines identifiées par spectrométrie de masse.

Sur la fraction nucléaire, en IP-Flag/HA, analyse par découpe de gel (GBA) et MudPIT, sur IP Flag et passage en gel Filtration (Pool A) par MudPIT, sur Extrait S100 et passage en gel Filtration (Pool A) par MudPIT. La protéine AID a été identifiée dans chaque fraction. Le faible taux de recouvrement entre les méthodes est sans doute dû à la complexité du mélange. Dans les cas du GBA, peu de protéines sont analysées à la fois, mais une perte importante est causée par la dilution des échantillons lors des préparations. Le MudPIT permet une séparation plus fine, mais avec un pool important de protéines. Par contre, réaliser une IP, puis une gel filtration et une analyse par MudPIT est très intéressante, car c'est ainsi que j'ai pu identifier les complexes PAF, Cohésine et Mediator.

5 - Co-immunoprécipitation des partenaires de AID

Après l'identification des protéines, il est impératif de confirmer l'interaction entre la nouvelle protéine identifiée (X) et AID par le biais d'une réaction de co-immunoprécipitation (IP). Cette étape cruciale nécessite une approche en deux temps (Figure 19). Tout d'abord, une IP Flag (AID) est réalisée pour corroborer le résultat de l'interaction d'AID avec la protéine X. Ensuite, une IP est effectuée avec l'anticorps spécifique de la protéine X, visant à immunoprécipiter la protéine AID. La figure ci-

dessous illustre la série d'immunoprécipitations que j'ai entreprises pour valider la découverte des divers complexes, tels que PAF, Cohésine et Mediator.

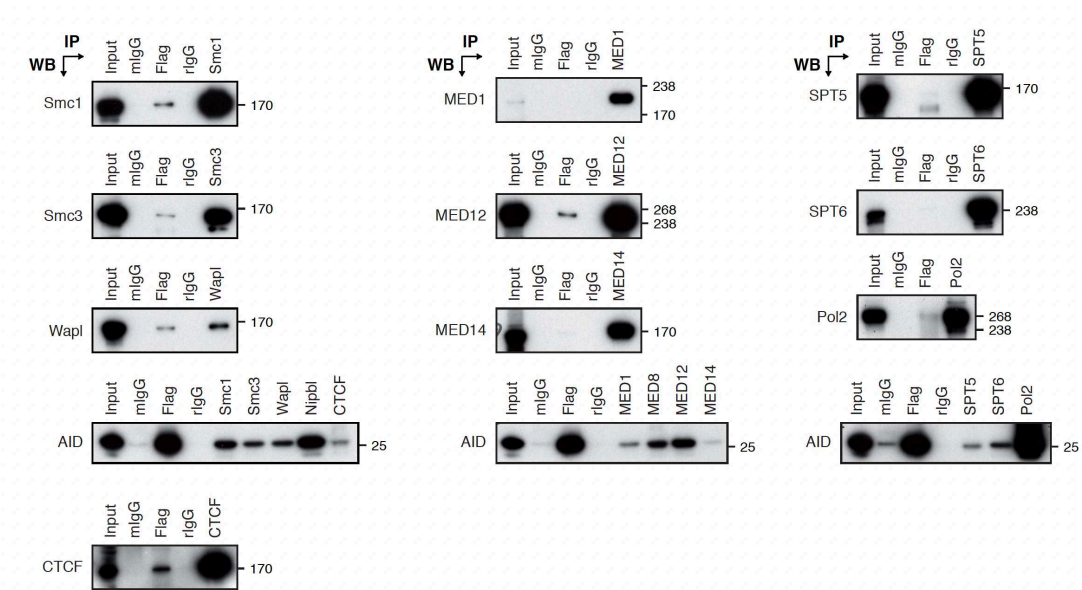


Figure 19: Co-immunoprécipitations des différents complexes associés avec AID.

Les extraits nucléaires ont été préparés à partir des cellules CH12 surexprimant AID-Flag-HA. Chaque immunoprécipitation (IP) a été effectuée en utilisant 1 mg d'extrait nucléaire. Les témoins négatifs ont été générés en utilisant un sérum correspondant à l'origine de l'anticorps.

6 - Le complexe Polymerase Associated Factor (PAF)

Par mes expériences j'ai identifié, en association avec AID, plusieurs sous-unités du complexe PAF (PAF1, CTR9 et Leo1). Ces résultats ont été confirmés par l'équipe de Svend Petersen-Mahrt (IFOM, Milan), qui a entrepris une approche protéomique similaire mais utilisant une autre lignée B et en purifiant la protéine AID endogène. En collaboration avec ce laboratoire, nous avons alors montré que le complexe PAF joue un rôle majeur dans la diversification des immunoglobulines lors du recrutement chromatinien d'AID (Willmann, 2012).

7 - Le complexe Cohésine

Le troisième complexe que j'ai découvert en association avec AID est le complexe cohésine, formé par les sous-unités Smc1, Smc3, Nipbl et Wapl, qui est impliqué dans la cohésion des chromatides, la réparation de l'ADN et dans la formation des boucles dans l'ADN, entre promoteurs et enhanceurs, en collaboration avec le complexe Mediator. Nous avons pu montrer que le complexe cohésine contrôle la commutation

isotypique et que la sous-expression conduit à un défaut de commutation isotypique et de réparation de l'ADN (Thomas-Claudepierre, 2013).

8 - Le Mediator

Etant donné que le complexe Mediator semble impliqué dans la formation des boucles dans l'ADN, qui sont nécessaires au mécanisme de commutation isotypique, nous avons également étudié son implication. Par des expériences de knockdown, et en utilisant un modèle murin d'inactivation conditionnelle dans les cellules B de la sous-unité Med1 du complexe Mediator, nous avons pu montrer que ce complexe est également nécessaire pour la commutation isotypique (Thomas-Claudepierre, 2016).

En conclusion de cette partie, ma recherche de co-facteurs régulant AID a permis la mise en évidence de nombreux complexes protéiques, et plus de 300 protéines comprenant des facteurs de transcription, des protéines d'organisations de la chromatine et des protéines de réparation de l'ADN. Ces découvertes ont abouti à de nombreuses publications, au niveau de l'équipe ou en collaboration et à l'obtention de plusieurs financements. Les 4 principales publications sont en annexe.

2 - Le Système d'Expression Universel (UES)

1 - Introduction

Le système UES, un nouveau système de clonage modulaire. L'étude fonctionnelle des protéines potentiellement régulatrices de AID nécessite la création de nombreux vecteurs d'expression, de recombinaison homologue et/ou de knockout. Les méthodes traditionnelles de clonage dans les vecteurs d'expression existants, dépendent des enzymes de restriction à site unique (qui ne sont souvent pas compatibles). De plus, elles offrent un choix très limité dans les rapporteurs fluorescents, les marqueurs de sélection et peuvent s'avérer très longues à réaliser. Pour contourner ces limitations, j'ai développé le système d'expression universel qui permet la production de vecteurs d'expression dans une seule réaction de clonage et en utilisant une enzyme de restriction unique. Ce système est composé de vecteurs de destination et d'un ensemble de modules. Ce système utilise la technologie de clonage "Golden Gate", qui repose sur l'utilisation d'enzymes de restriction de type IIS (Figure 20). Ces enzymes coupent à distance de leur séquence cible, et génèrent des extrémités qui peuvent être manipulées de manière à dicter l'orientation de clonage et la compatibilité entre différents fragments d'ADN. Tel un Légo, jusqu'à 8 fragments d'ADN peuvent être ainsi assemblés dans un ordre bien précis (Figure 20).

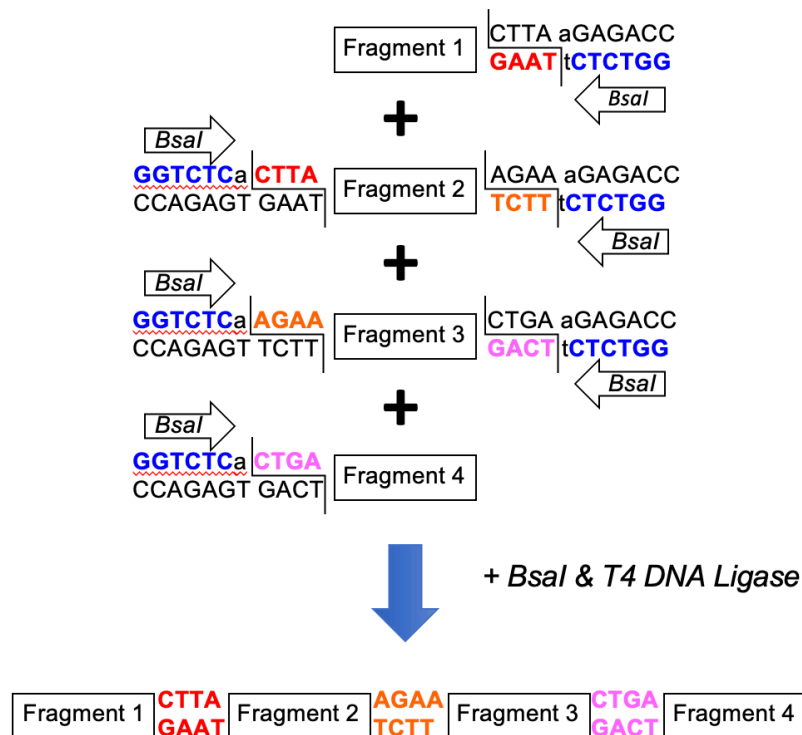


Figure 20: Digestion et ligation simultanée des fragments d'ADN.

La digestion par l'enzyme de restriction *BsaI* laisse des extrémités 3' débordantes permettant de dicter l'orientation du clonage et le positionnement des fragments d'ADN. Notez que les sites de restriction *BsaI* (indiqués en gras) sont perdus après la ligation des extrémités compatibles (en couleur), ce qui permet la digestion et la ligation en simultanée.

2 - Principe

Le Système d'Expression Universel utilise le principe de digestion/ligation en simultanée pour l'assemblage des vecteurs d'expression (Figure 21). Un vecteur de destination donnée (Figure 22) est mélangé avec une sélection d'éléments de la librairie de modules (Figure 23) ou de produits de PCR, dont les sites de restriction *BsaI* ont été intégrés lors de l'amplification. L'ensemble est incubé simultanément avec l'enzyme *BsaI* et l'ADN ligase 4, transformé dans la bactérie *E. Coli*, puis étalé sur une boîte de culture contenant l'antibiotique approprié de manière à sélectionner uniquement le vecteur de destination. Les vecteurs de destination contiennent une cassette LacZ, qui est perdue lors de l'assemblage, et permet un criblage blanc/bleu. Seuls les vecteurs de destination ayant incorporé les modules pousseront et donneront des colonies blanches (Figure 21). Quelques clones positifs seront mis en culture pour réaliser des préparations d'ADN plasmidiques qui seront contrôlées par digestion enzymatique et séquençage. Seulement 4 jours sont nécessaires pour réaliser cet assemblage.

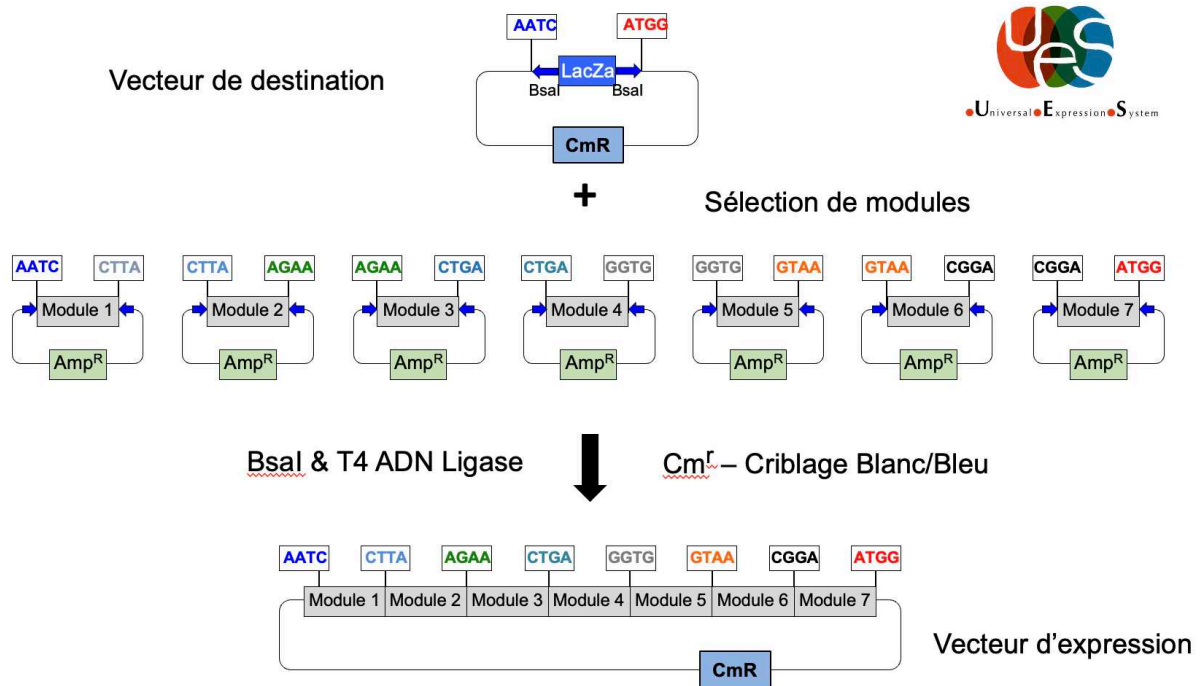


Figure 21: Vue d'ensemble de la procédure d'assemblage du vecteur d'expression.

Les vecteurs de destination possèdent un système de sélection bleu/blanc, une origine de réplication bactérienne et un gène de résistance au chloramphénicol pour la culture et la sélection.

3 - Les vecteurs de destination

Pour commencer, j'ai construit plusieurs vecteurs de destination (Figure 22). Un premier vecteur (pX) ne contient pas d'éléments transcriptionnels et me permet d'assembler des vecteurs destinés à la recombinaison homologe en vue de la génération de souris transgéniques, par exemple. Un deuxième vecteur (pCXpA) contient un promoteur CMV et un signal de poly-adylation, permettant l'expression des cDNAs (modules) par le promoteur CMV. Un troisième (pXPA) contient uniquement un signal de poly-adylation et permet le choix du promoteur (inductible, constitutif, fort, tissu-spécifique, etc). Finalement, les modules assemblés dans ces vecteurs peuvent être transférés dans un vecteur rétroviral (pRX), de manière à pouvoir infecter des cellules qui sont difficilement transfectables.

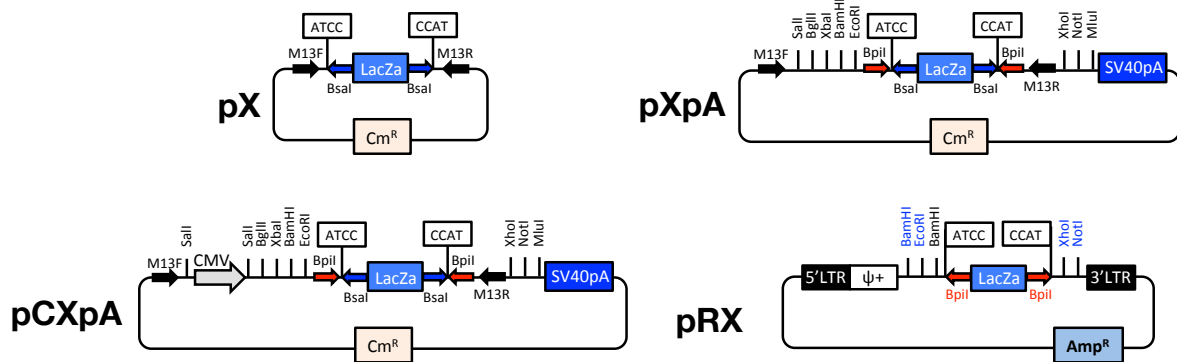


Figure 22: Les vecteurs de destinations.

4 - Les modules

Ma librairie de modules (Figure 23) est constituée de plasmides qui contiennent l'ensemble des éléments nécessaires à l'élaboration des vecteurs d'expression. J'ai construit ces modules par PCR. Ils contiennent des sites de restriction Bsal qui dictent leur position de clonage (de 1 à 7). Les modules contiennent différents éléments: promoteurs, protéines fluorescentes, épitopes synthétiques, cassettes de sélection, etc. La combinaison de vecteurs de destination et la librairie de modules ($n > 200$ actuellement) me permet donc de concevoir et réaliser un ensemble de vecteurs d'expression très varié. Notez que le système n'est pas limité à l'assemblage de 7 modules et que je peux générer des vecteurs d'expression contenant moins de modules, en choisissant des extrémités adaptées.

| Module 1 | Module 2 | Module 3 | Module 4 | Module 5 | Module 6 | Module 7 |
|-------------------|--------------------------------|---------------------|--------------------------------|---------------------|--------------------------------|-------------------------------|
| ATCC-N-TATG | TATG-N-GAAC | GAAC-N-CGCA | CGCA-N-CAGG | CAGG-N-AGCA | AGCA-N-GCTT | GCTT-N-CCAT |
| Promoteurs | | Bi-cistronic | | | Protéines Fluorescentes | |
| CMV | Protéines Fluorescentes | IRES | Protéines Fluorescentes | Promoteurs | EGFP | Cassettes de sélection |
| SV40 | EGFP | T2A | EGFP | CMV | EYFP | SV40-Puro ^R |
| PGK | EYFP | | EYFP | SV40 | ECFP | SV40-Neo ^R |
| TRE | ECFP | Epitopes | ECFP | PGK | mCherry | SV40-Hygro ^R |
| | mCherry | Flag | mCherry | UbC | paEGFP | SV40-Zeo ^R |
| Knockout | sEYFP ₁₋₁₅₄ | HA | | | Dendra2 | |
| U6-gRNA | sEYFP ₁₅₅₋₂₃₈ | Flag-HA | Epitopes | Bi-cistronic | | |
| GLN1-gRNA | | BioTag | Flag | IRES | Epitopes | |
| | Epitopes | | HA | T2A | Flag | |
| | | Knockout | Flag-HA | | HA | |
| | Flag | U6-gRNA | BioTag | Epitopes | Flag-HA | |
| | HA | GLN1-gRNA | | Flag | Myc | |
| | Flag-HA | | Inducible | HA | V5 | |
| | BioTag | | Tet-ON | Myc | BioTag | |
| | | | Tet-OFF | V5 | | |
| | Knockout | | | BioTag | | |
| | U6-gRNA | | Knockdown | | | |
| | GLN1-gRNA | | U6-shRNA | CRISPR-Cas9 | | |
| | | | | Cas9-HF | | |
| | | | | nCas9-HF | | |
| | | | | dCas9-HF | | |

Figure 23: Liste non exhaustive des modules disponibles.

5 - Le système CRISPR/Cas9

Le système CRISPR/Cas9 est un système composé d'un ou plusieurs guides ANR (20 nucléotides) qui se fixent sur l'ADN et qui sont reconnus par la nucléase Cas9. Le système CRISPR/Cas9 conduit à la génération de cassures double brin à l'ADN, de knockouts ou de knockins (Figure 24).

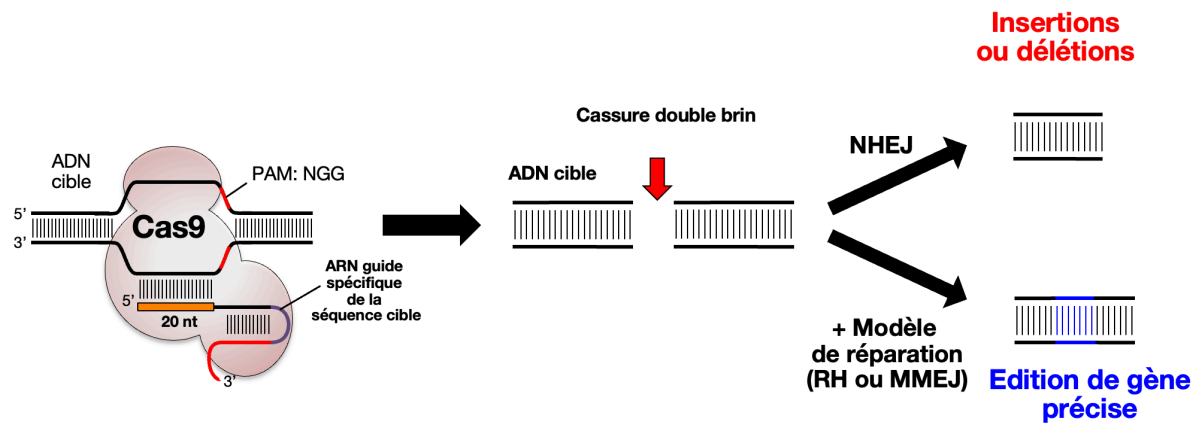


Figure 24: Le système CRISPR-Cas9.

Utilisation de 2 voies de réparations pour réaliser le knock-in, par la voie de la recombinaison homologue (RH) ou par la voie de la micro-homologie (MMEJ). La voie de la Jonction d'Extrémités Non Homologues (NHEJ en anglais) permet des insertions ou délétions dans le génome.

J'ai standardisé le système UES pour obtenir une efficacité optimale dans la conception de vecteurs contenant deux guides ARN (Figure 25), construction qui est la plus adaptée à nos expérimentations pour obtenir des knockouts.

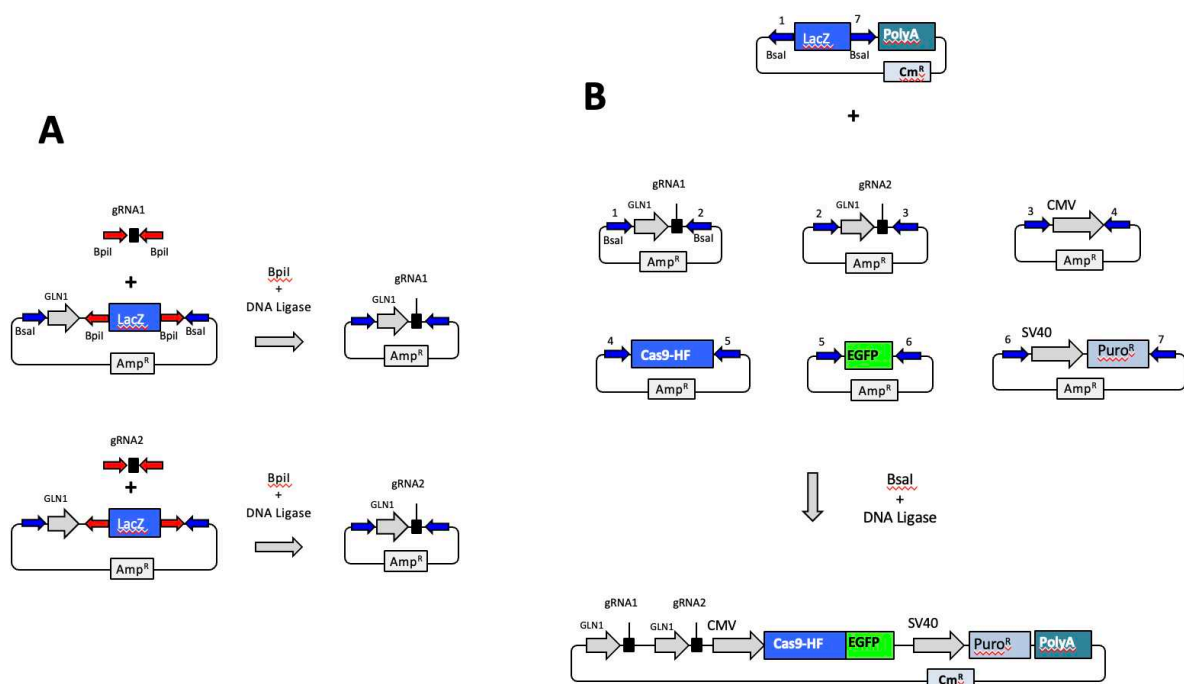


Figure 25: Création d'un vecteur CRISPR-Cas9 avec deux guides ARN.

Les deux gARN sont clonés par digestion/ligation. Puis ils sont intégrés dans le vecteur final, qui contient l'ensemble des éléments nécessaires à la conception d'un knockout.

J'ai adapté le système UES (Figure 26) afin de générer des vecteurs complexes pouvant en effet contenir jusqu'à 3 guides ARN ainsi que la Cas9 fusionnée à des

rapporteurs fluorescents (GFP, YFP, CFP, etc.) et à des cassettes de sélection (Puromycine, Néomycine, Hygromycine, etc.).

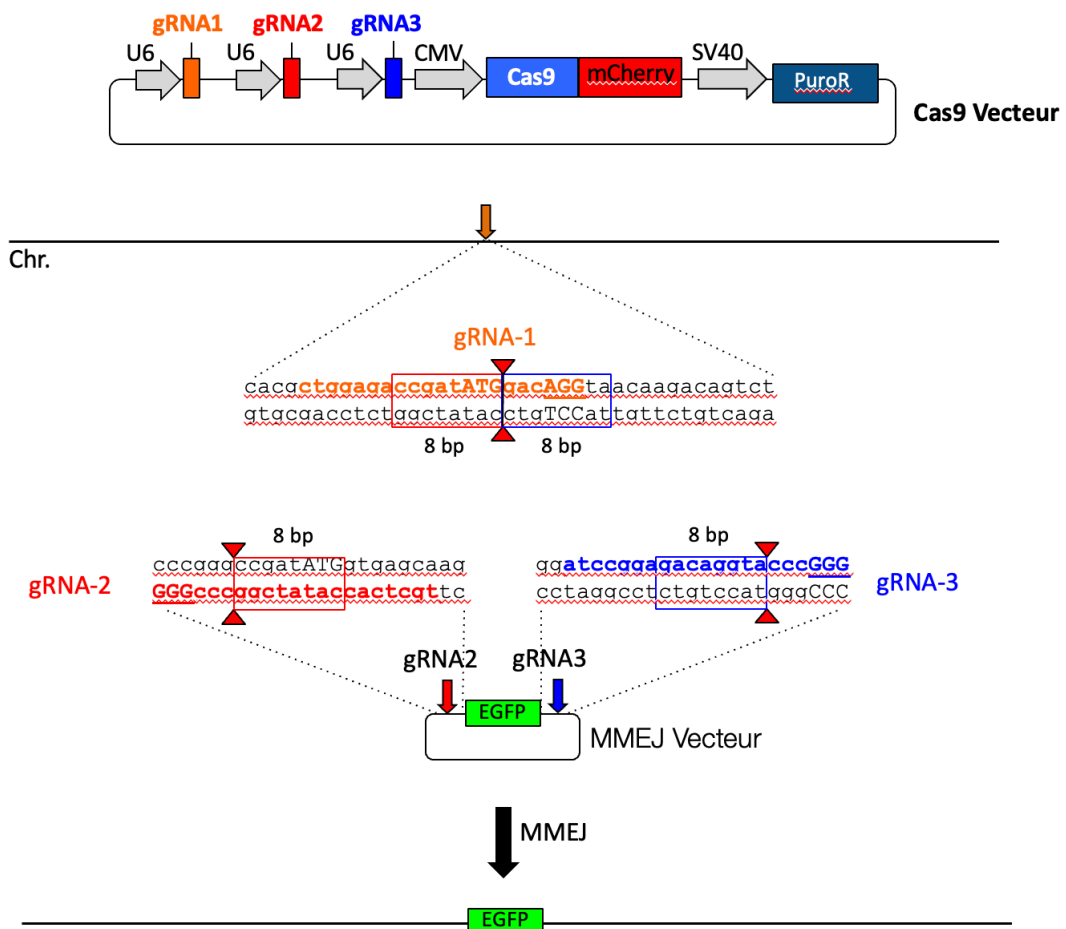


Figure 26: Knock-in par microhomologie (MMEJ).

Le vecteur le plus utilisé dans notre laboratoire pour créer un knock-out est sans conteste le vecteur Crispr-Cas9 avec deux guides ARN intégrés.

6 - Développement du système UES

Je forme et accompagne aux méthodes de clonages complexe nombre d'étudiants de l'IGBMC ou d'autres instituts ainsi qu'aux ingénieurs de l'École Supérieure de Biotechnologie de Strasbourg (ESBS). Grâce au système UES et ses multiples possibilités d'évolution, j'ai collaboré avec un nombre important d'équipes au sein de l'IGBMC et d'ailleurs, et 5 articles ont été publiés en co-auteur (Lemaître C, 2014; Tsouroula K, 2016; Soria-Bretones I, 2017 ; Beck C, 2018; Pizon V, 2020). J'ai ainsi créé des vecteurs permettant la visualisation de locus unique dans le génome en utilisant la version dCas9 (Figure 27 et 28).

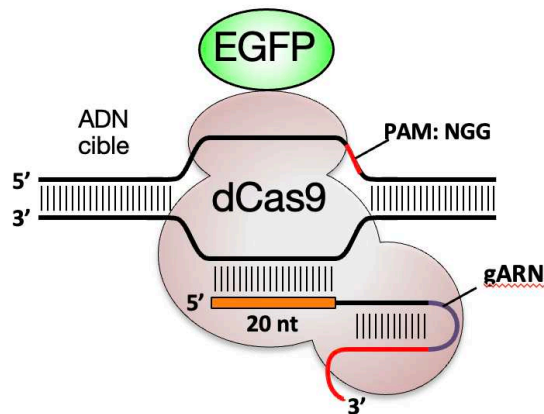


Figure 27: La dCas9.

La dCas9 (qui se fixe mais ne coupe pas) est fusionnée à une protéine fluorescente (EGFP). Elle peut être ciblée sur n'importe quelle localisation génomique.

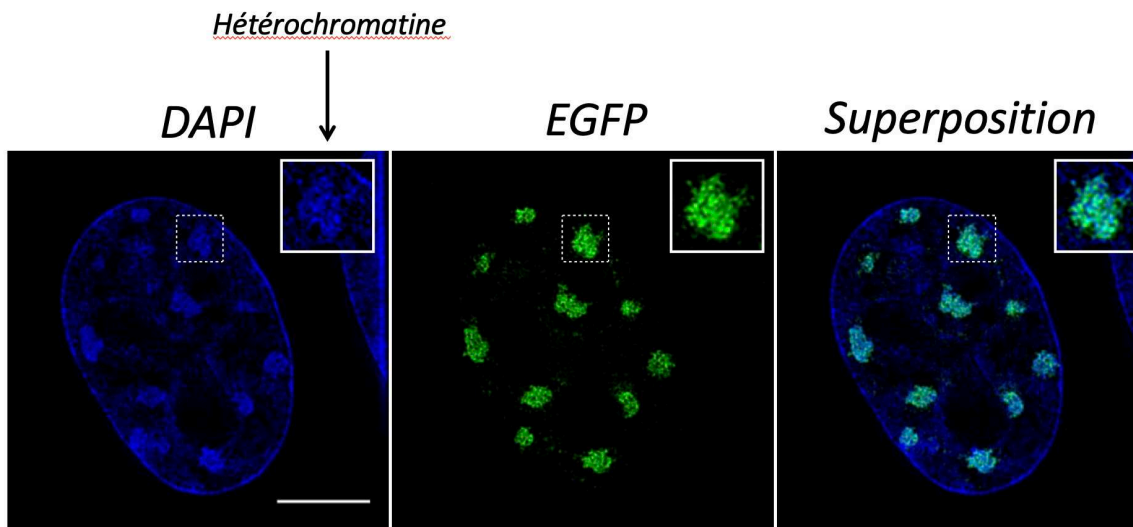


Figure 28: Visualisation de l'hétérochromatine avec CRISPR.

Les cellules exprimant la dCas9 et un gRNA spécifique des régions de satellites majeurs (régions péricentromériques). Le DAPI est un fluorochrome qui se complexe sélectivement avec l'ADN double brin.

Afin de faire bénéficier du système UES à l'ensemble de la communauté scientifique, j'ai supervisé le transfert de cette technologie vers le service de Biologie Moléculaire de l'IGBMC, dont j'ai formé le personnel. J'ai également participé à la création d'un site web dédié par le service informatique de l'IGBMC, qui permet aux chercheurs (Figure 29) de commander directement en ligne les vecteurs d'expression qu'ils désirent (jusqu'à 100/an; ues.igbmc.u-strasbg.fr).

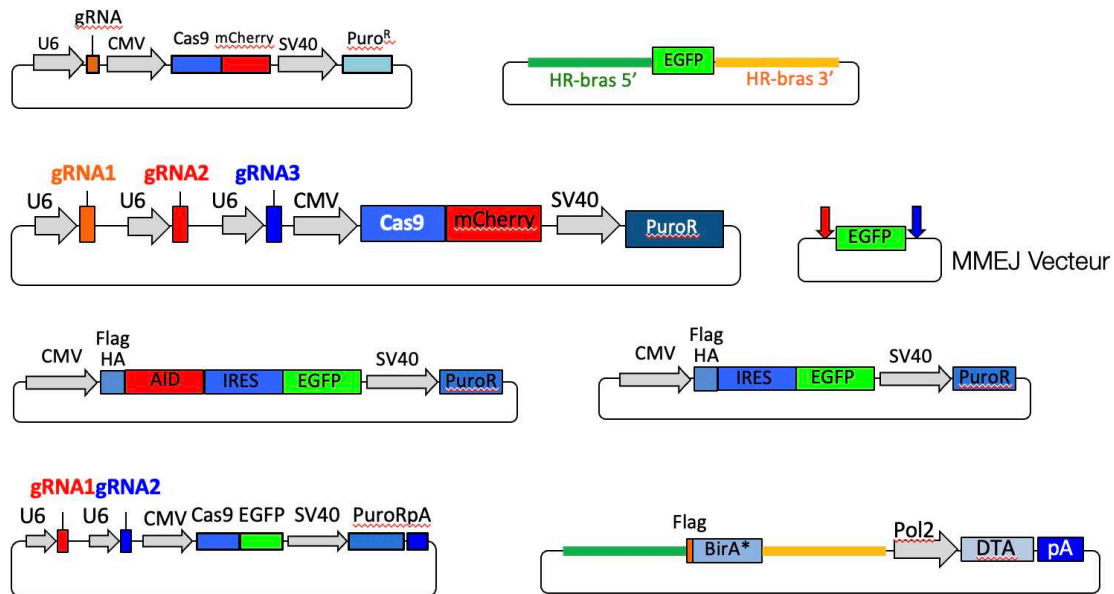


Figure 29: UES. Une gamme non exhaustive.

7 - Un criblage à l'échelle du génome

Un criblage CRISPR-Cas9 pour identifier les gènes impliqués dans la CI.

Le lancement de cette étude nécessite la conception de vecteurs plasmidiques permettant l'expression de la protéine Cas9 dans la lignée CH12, tout en intégrant un marqueur fluorescent (mCherry) et une résistance à un antibiotique (l'hygromycine). Après l'obtention d'une lignée stable, j'ai développé une série de vecteurs facilitant l'utilisation des bibliothèques de gRNA nécessaires à la réalisation de ce criblage (Figure 30).

Pour identifier les gènes impliqués dans la CI, nous avons conduit un crible CRISPR/Cas9 à l'échelle du génome en utilisant la lignée CH12 et la bibliothèque de gRNAs mBrie (80 000 gRNAs, 4 gRNAs/gène ; 20 000 gènes). Nous avons ainsi identifié FAM72A (Family with Sequence Similarity 72, Member A), une protéine qui est surexprimée dans plusieurs cancers et qui interagit avec l'isoforme nucléaire (UNG2) de l'enzyme Uracile DNA Glycosylase, qui enlève les uraciles générés par

AID. Nous avons montré que l'interaction FAM72A-UNG2 contrôle les niveaux d'UNG2 et que ceci est important car l'absence de FAM72A, dans les cellules CH12 ou *in vivo* chez la souris, cela conduit à un défaut de CI et de HMS. Par ailleurs, nous avons clairement démontré que ce phénotype est dû à l'augmentation du niveau protéique d'UNG2.

Nous avons alors proposé un modèle dans lequel FAM72A interagit avec UNG2 pour contrôler son niveau physiologique en déclenchant sa dégradation, régulant le niveau d'excision de l'uracile et donc l'équilibre entre la réparation fidèle ou non-fidèle de l'ADN. Nos résultats ont des implications potentielles pour la tumorigenèse, car des niveaux réduits d'UNG2 médiés par la surexpression de Fam72a déplaceraient l'équilibre vers la réparation mutagène de l'ADN, rendant les cellules plus susceptibles à acquérir des mutations. Ces résultats ont été publiés dans la revue **Nature** (Rogier 2021).

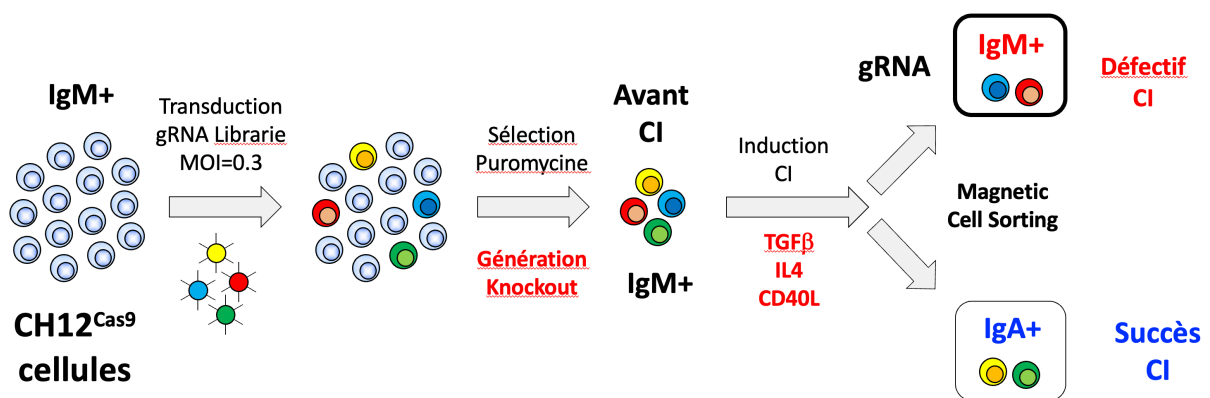


Figure 30: Le crible CRISPR/Cas9 à l'échelle du génome.

Le crible CRISPR/Cas9 à l'échelle du génome, en utilisant la librairie mBRIE, qui contient 19 000 genes, 78 000 gRNAs (4 gRNAs/gene). Le séquençage des gRNA dans la population IgM+ permettra d'identifier les gènes requis pour la CI.

8 - Perspectives

Les vecteurs d'expression les plus largement adoptés dans mes investigations sont les vecteurs CRISPR-Cas9 équipés de deux guides ARN. Ces vecteurs sont minutieusement optimisés pour permettre la conception précise de knock-out ou de knock-in, en fonction de la disponibilité du modèle de réparation. Cette technologie constitue un pilier essentiel de mes travaux de recherche, m'offrant une flexibilité inégalée dans l'étude du complexe protéique HIF (Hypoxia-Inducible Factor). Son utilisation constante se révèle être un atout majeur dans l'approfondissement de ma

compréhension des mécanismes moléculaires régissant la réponse cellulaire à l'hypoxie, contribuant ainsi significativement à l'avancement de mes recherches scientifiques.

3 - Étude du complexe protéique HIF

Il a été démontré que, au cours des réponses immunitaires, des zones d'hypoxie se développent au sein des centres germinaux (GC), influençant ainsi la réponse humorale. Toutefois, les mécanismes physiologiques qui régissent ces processus dans le microenvironnement des GC restent largement inconnus. Il est plausible que la tension d'oxygène et la réponse physiologique à l'hypoxie jouent un rôle essentiel dans la régulation de la réponse humorale au sein des GC. Selon ma conjecture, HIF1 pourrait jouer un rôle central dans la régulation de l'activation transcriptionnelle du gène AID au sein des cellules B, participant ainsi activement au processus de commutation isotypique (CI) au cours des réponses immunitaires. La question cruciale soulevée par cette conjecture est de savoir si Hif1a est indispensable pour une commutation isotypique optimale au sein des cellules CH12. Cette interrogation ouvre une voie intéressante pour explorer le lien entre la régulation de l'oxygène, la réponse à l'hypoxie, et les mécanismes cellulaires spécifiques qui coordonnent la commutation isotypique dans le contexte des réponses immunitaires.

1 - Hif1a est nécessaire pour une CI optimal dans les cellules CH12

Pour tester cette hypothèse, j'ai utilisé la technologie CRISPR-Cas9 pour inactiver le gène *Hif1a* (Figure 31) dans les cellules CH12, une lignée de cellules B murines. Cette lignée, lorsque cultivée in vitro en présence de TGF β , d'IL4 et de ligand CD40, est connue pour stimuler efficacement la commutation isotypique d'IgM en IgA. Après avoir obtenu un clone avec le gène *Hif1a* délété, confirmé par PCR, j'ai validé l'absence de la protéine HIF-1 α par Western Blot.

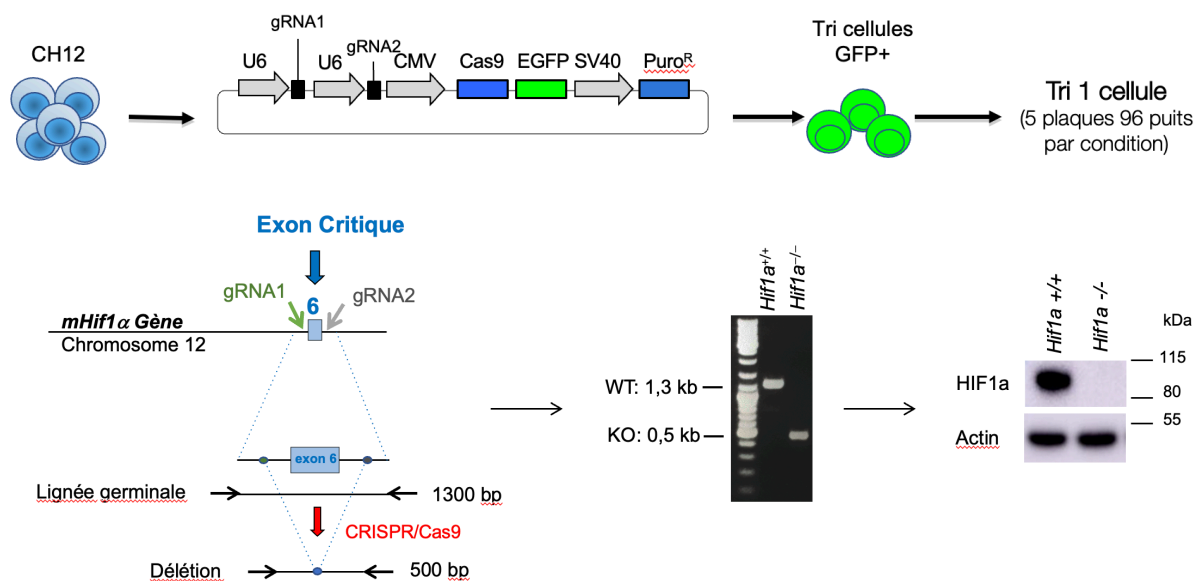


Figure 31: Élaboration du knock-out du gène Hif1a.

Les deux gRNA seront dirigés vers l'exon critique (6) du gène Hif1a. Les cellules seront ensuite triées par FACS (ARIA BD). Après une période de deux semaines, les clones seront soumis à une analyse par PCR afin de vérifier la suppression des deux allèles du gène. La confirmation de l'absence de la protéine HIF1α sera réalisée par PCR, puis par Western Blot.

Pour déterminer si la CI est affectée par l'inactivation de *Hif1a*, des cellules de type sauvage et *Hif1a*^{-/-} ont été stimulées (Figure 32). Après 72 heures, l'efficacité de la CI a été évaluée par cytométrie en flux. Alors que la lignée cellulaire parentale présentait 60 % de cellules exprimant des IgA, ce pourcentage est réduit de 2,5 fois dans les cellules CH12 *Hif1a*^{-/-}.

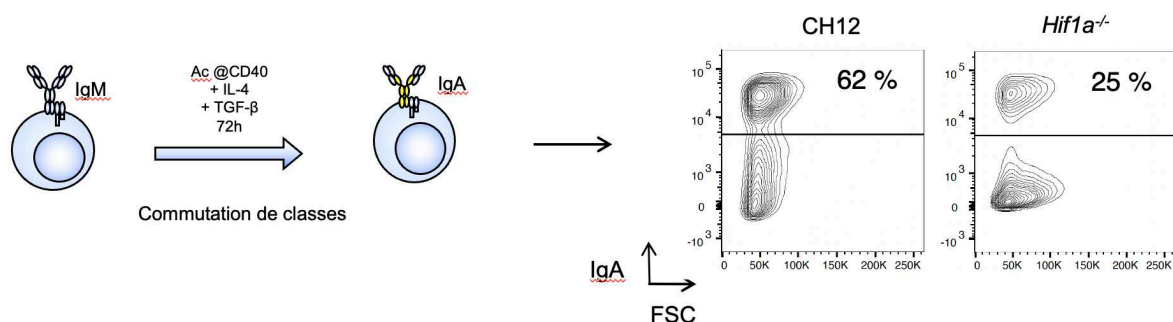


Figure 32: Détermination de l'efficacité du knock-out du gène Hif1a.

Après 72 heures, l'efficacité de la CI a été évaluée en déterminant le pourcentage de cellules IgA⁺ par cytométrie de flux. Ce pourcentage est passé de 62% à 25% dans le clone *Hif1a*^{-/-}.

Comme la CI dépend de la prolifération (Figure 33), j'ai induit la CI dans des cellules marquées au colorant CFSE (carboxyfluorescéine diacétate succinimidyl ester) et j'ai constaté que tous les génotypes diluaient le CFSE de la même façon.

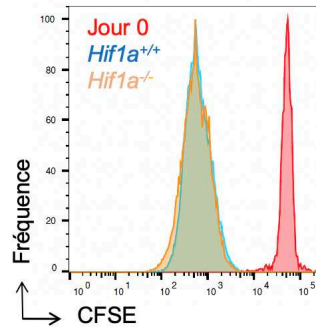


Figure 33: Marquage des cellules stimulées au CFSE.

Le colorant CFSE est introduit dans la cellule, puis va se diluer au fur et à mesure de la prolifération des cellules. En rouge nous avons l'échantillon avec la concentration de référence de départ. Après prolifération, nous avons en bleu le clone sauvage, en jaune les cellules *Hif1a*^{-/-}. La superposition des deux courbes montre une absence de défaut de prolifération.

Pour confirmer que le phénotype CI défectueux observé est dû à l'absence de *Hif1a* (Figure 34), j'ai transduit les cellules *Hif1a*^{-/-} avec un rétrovirus exprimant HIF-1 α . J'ai vérifié par Western Blot que la protéine HIF-1 α est bien exprimée (Figure 35).

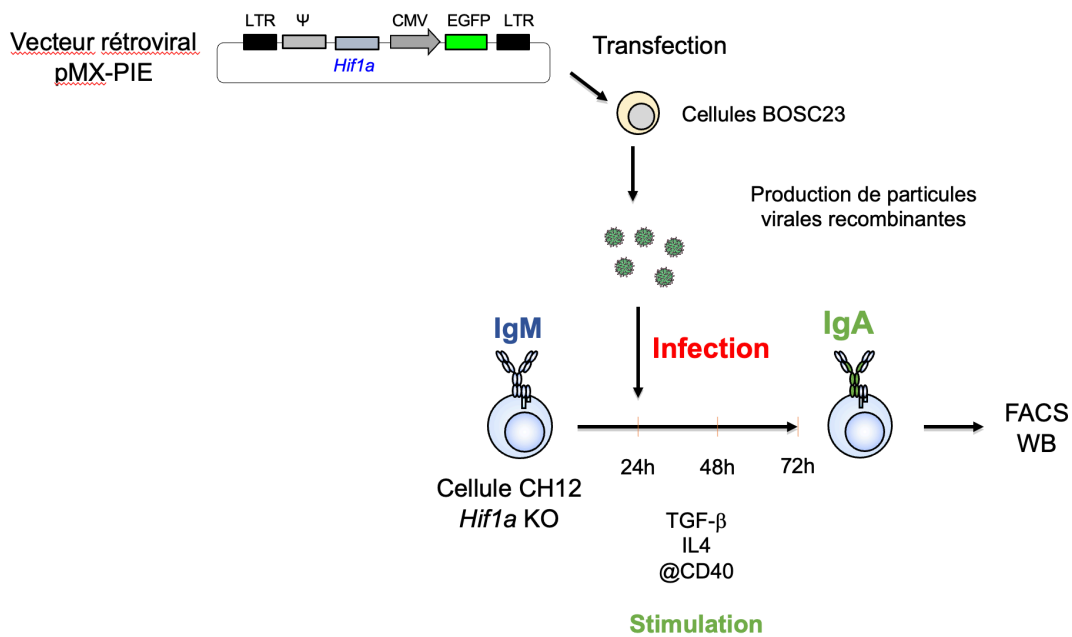


Figure 34: Infection rétrovirale avec pMX-PIE.

Le rétrovirus permettant le ré-expression de la protéine HIF1 α est transfecté dans des cellules Bosc23 qui permettent la production de particules virales qui infecteront les cellules CH12 sans le gène *Hif1a*. Après mise en culture et 72h de stimulation, un western blot permet de confirmer l'expression de la protéine sur-exprimée, et un FACS est réalisé pour déterminer le % d'IgA.

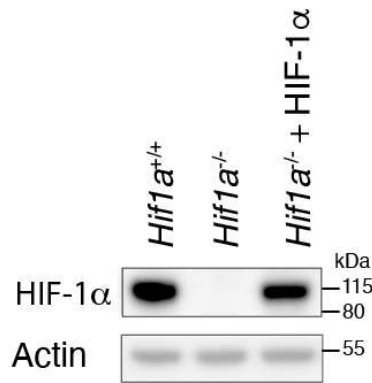


Figure 35: Western Blot de la reconstitution de HIF-1 α .

Western blot pour détecter la protéine HIF-1 α , et b-Actin dans les cellules CH12 de type sauvage, les cellules *Hif1a*^{-/-}, et les cellules *Hif1a*^{-/-} ré-exprimant la protéine HIF-1 α . Les poids moléculaires en kDa sont indiqués.

J'ai montré que la réexpression de HIF-1 α est capable de restaurer de manière significative le pourcentage de cellules exprimant IgA (Figure 36).

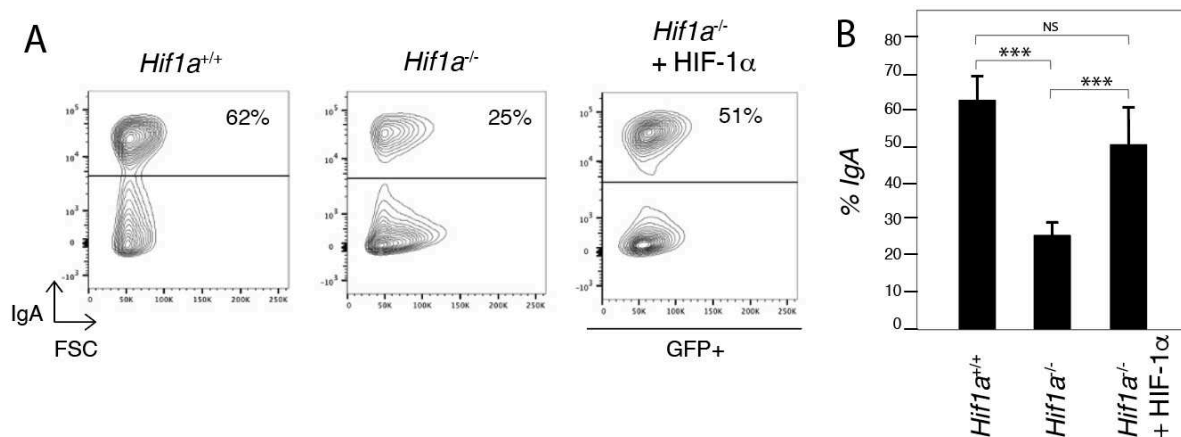


Figure 36: Restauration du phénotype *Hif1a*^{-/-} (FACS).

(A) Analyse par cytométrie en flux de l'expression des IgA dans les cellules sauvages, *Hif1a*^{-/-} et *Hif1a*^{-/-} ré-exprimant HIF-1 α . Les graphiques représentatifs sont présentés. (B) Pourcentage moyen de cellules IgA⁺ provenant de cinq expériences indépendantes dans les cellules sauvages, *Hif1a*^{-/-} et *Hif1a*^{-/-} ré-exprimant HIF-1 α . Valeur p (test t de Student bilatéral; ***p < 0.001, NS p > 0,05); les valeurs p sont *Hif1a*^{+/+} contre *Hif1a*^{-/-} = 5,05 × 10⁻⁶, les valeurs p sont *Hif1a*^{-/-} contre *Hif1a*^{-/-} + HIF-1 α = 0,00066, les valeurs p sont *Hif1a*^{+/+} contre *Hif1a*^{-/-} + HIF-1 α = 0,05264.

Je conclus donc que *Hif1a* est nécessaire pour obtenir une CI optimale dans les cellules CH12, indépendamment des défauts de prolifération.

2 - L'induction de l'ARNm AID est compromise dans les cellules *Hif1a*^{-/-}

Pour déterminer si l'expression de AID est déficiente en l'absence de *Hif1a* (Figure 37), j'ai quantifié le niveau de protéine AID par Western blot et par RT-qPCR pour les ARNm.

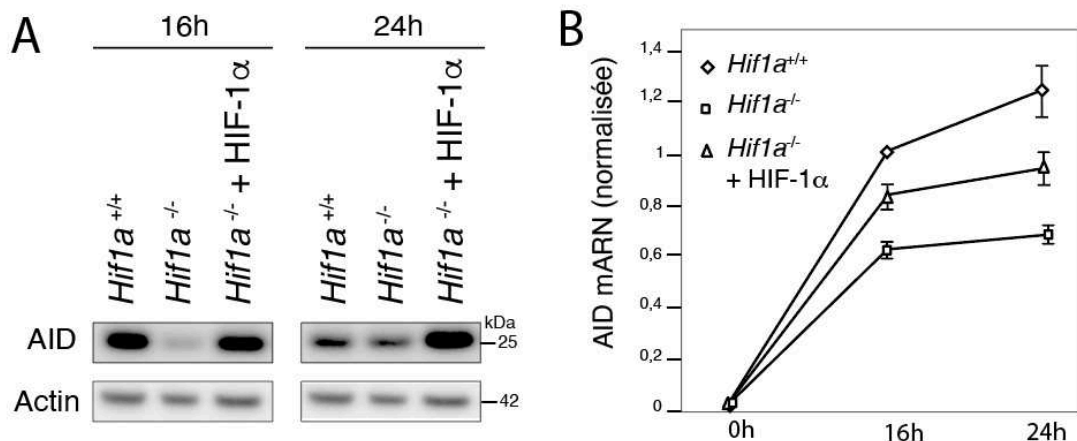


Figure 37: Restauration du phénotype *Hif1a*^{-/-} (ARN, protéine).

(A) Analyse par Western blot de l'AID et de la β -Actine dans les cellules sauvages, *Hif1a*^{-/-} et *Hif1a*^{-/-} réexprimant HIF-1 α . Les poids moléculaires (kDa) sont indiqués. (B) Histogramme montrant les niveaux de transcription de l'*Aicda* dans les cellules sauvages, *Hif1a*^{-/-} et *Hif1a*^{-/-} réexprimant HIF-1 α , déterminés par RT-qPCR. Les valeurs moyennes des triplicats techniques de trois expériences indépendantes ont été normalisées par rapport aux transcrits HPRT (méthode des CT) et sont indiquées par rapport aux cellules WT, qui ont été normalisées à un.

J'ai constaté que l'induction de l'ARNm de AID est compromise dans les cellules *Hif1a*^{-/-} et que cela entraîne un retard dans l'expression de AID au niveau protéique et un défaut concomitant dans la CI. De manière intéressante, j'ai pu observer que la réexpression de HIF-1 α peut restaurer les niveaux d'ARNm et de protéine AID.

3 - *Hif1b* est nécessaire pour l'expression optimale d'AID et de l'efficacité de la CI

La translocation nucléaire de HIF-1 α est dépendante de HIF-1 β . Ainsi, en l'absence de HIF-1 β , je devrais observer un phénotype similaire. Pour tester cette hypothèse, j'ai inactivé le gène *Hif1b* (Figure 38) dans les cellules CH12. L'absence de HIF-1 β a été vérifiée par Western blot.

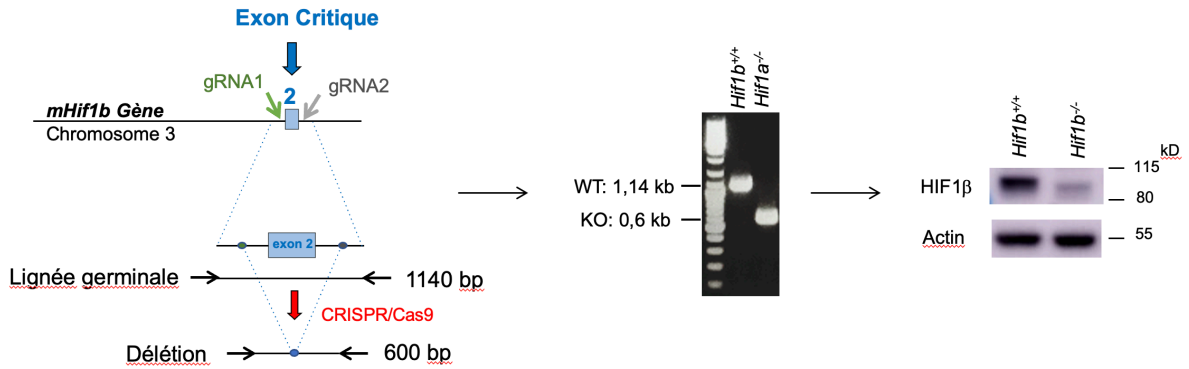


Figure 38: Élaboration du knock-out du gène *Hif1b*.

Les deux gRNA vont cibler l'exon critique (2) du gène *Hif1b*. Les cellules seront triées par FACS (ARIA BD). Après 2 semaines, les clones seront analysés par PCR afin de déterminer si les deux allèles du gène ont bien été délévés. L'absence de la protéine HIF-1β est vérifiée par western blot.

Pour déterminer si la CI est affectée par l'inactivation de *Hif1b*, des cellules de type sauvage ou *Hif1b*^{-/-} ont été stimulées. Alors que la lignée cellulaire parentale présentait 55 % de cellules exprimant l'IgA, ce pourcentage est significativement réduit dans les cellules *Hif1b*^{-/-}. Comme les cellules *Hif1a*^{-/-}, les cellules *Hif1b*^{-/-} ont proliféré au même niveau que les cellules de type sauvage (Figure 39).

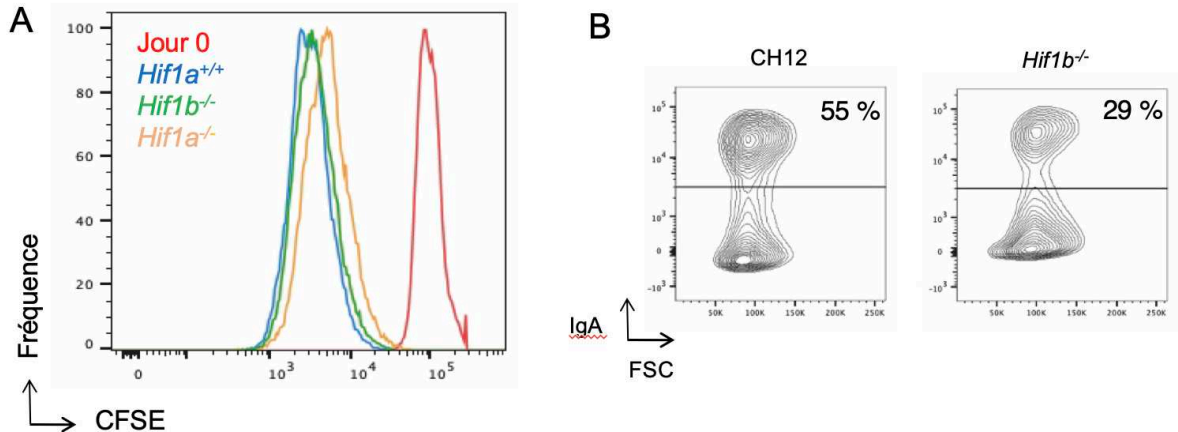


Figure 39: Marquage des cellules stimulées au CFSE.

(A) Marquage des cellules stimulées au CFSE. En rouge le temps de référence de départ, en bleu le clone sauvage, en jaune les cellules *Hif1a*^{-/-}, en vert les cellules *Hif1b*^{-/-}. La superposition des trois courbes montre une absence de défaut de prolifération. (B) Après 72 heures, l'efficacité de la CI a été évaluée en déterminant le pourcentage de cellules IgA⁺ par cytométrie de flux. Ce pourcentage est passé de 55% à 29% dans le clone *Hif1a*^{-/-}, indiquant une baisse significative de CI.

Pour confirmer que le phénotype défectueux de CI observé est dû à l'absence de HIF-1b, j'ai transduit les cellules *Hif1b*^{-/-} avec un rétrovirus exprimant HIF-1b (Figure 40). Je constate que la réexpression de HIF-1b est capable de corriger le défaut de CI.

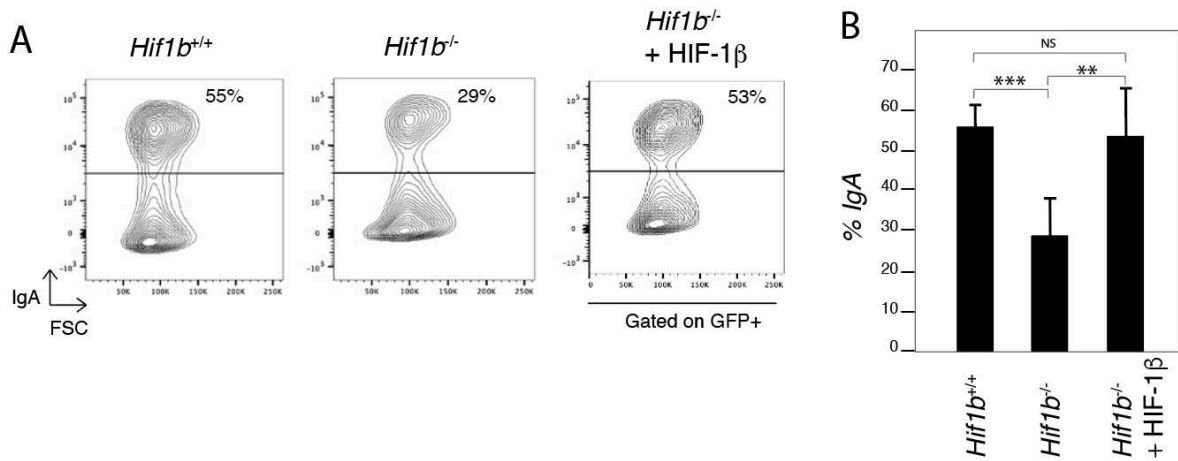


Figure 40: Restauration du phénotype *Hif1b*^{-/-} (FACS).

(A) Analyse par cytométrie en flux de l'expression des IgA dans les cellules sauvages, *Hif1b*^{-/-} et *Hif1b*^{-/-} réexprimant HIF-1b. Les graphiques représentatifs sont présentés. (B) Pourcentage moyen de cellules IgA⁺ provenant de cinq expériences indépendantes dans les cellules sauvages, *Hif1b*^{-/-} et *Hif1b*^{-/-} réexprimant HIF-1b. Valeur p (test t de Student bilatéral; ***p < 0.001, NS p > 0,05); les valeurs p sont *Hif1b*^{+/+} contre *Hif1b*^{-/-} = 0,0007, les valeurs p sont *Hif1b*^{-/-} contre *Hif1b*^{-/-} + HIF-1α = 0,0079, les valeurs p sont *Hif1b*^{+/+} contre *Hif1b*^{-/-} + HIF-1α = 0,75732.

Tout comme la déficience en *Hif1a*, les cellules *Hif1b*^{-/-} présentent un retard dans l'expression de l'ARNm et des niveaux de protéines de AID (Figure 41). Je conclus donc que *Hif1b* est nécessaire pour une expression optimale de AID et de CI dans les cellules CH12, indépendamment des défauts de prolifération.

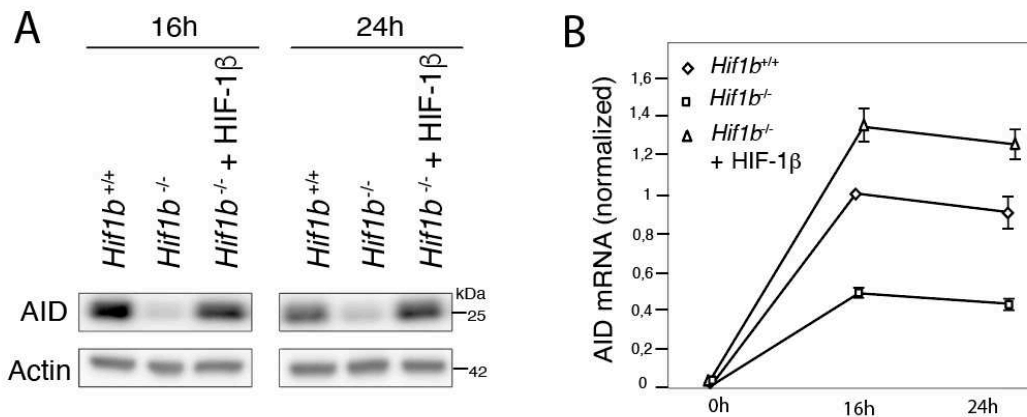
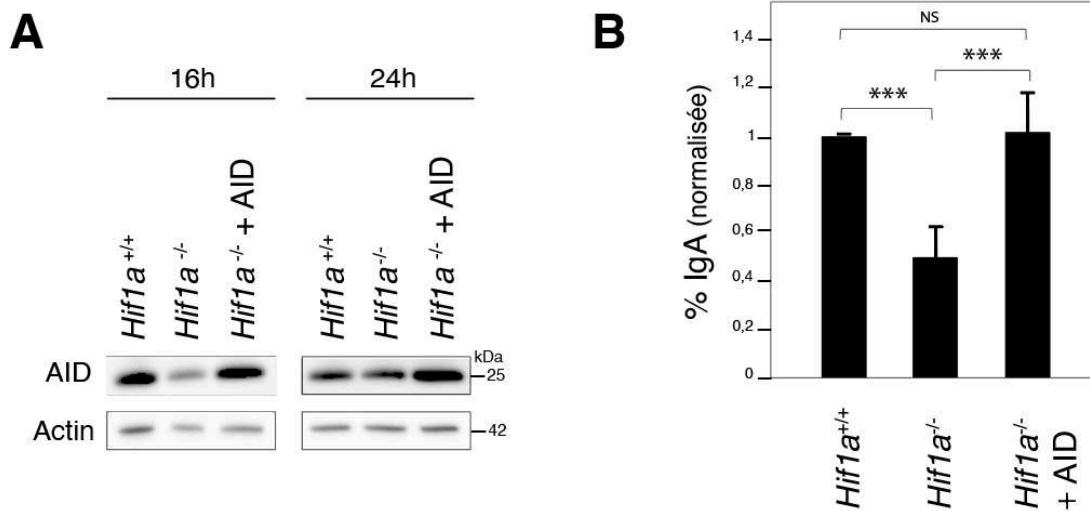


Figure 41: Restauration du phénotype *Hif1b*^{-/-} (ARN, protéine).

(A) Analyse par Western blot de l'AID et de la β-Actine dans les cellules sauvages, *Hif1b*^{-/-} et *Hif1b*^{-/-} réexprimant HIF-1β. Les poids moléculaires (kDa) sont indiqués. (B) Histogramme montrant les niveaux de transcription de l'*Aicda* dans les cellules sauvages, *Hif1b*^{-/-} et *Hif1b*^{-/-} réexprimant HIF-1β, déterminés par RT-qPCR. Les valeurs moyennes des triplicats techniques de trois expériences indépendantes ont été normalisées par rapport aux transcrits HPRT (méthode des CT) et sont indiquées par rapport aux cellules WT, qui ont été normalisé à un.

4 - Le défaut de CI dans les cellules B déficientes en Hif1a ou en Hif1b est dû à l'expression sous-optimale d'AID

Si HIF-1a et HIF-1b sont impliquées dans la médiation des niveaux optimaux de production d'ARNm de AID, alors le défaut de CI observé dans les cellules *Hif1a*^{-/-} et *Hif1b*^{-/-} devrait être résolu par la surexpression de la protéine AID (Figure 42). Pour vérifier cette hypothèse, j'ai transduit les cellules *Hif1a*^{-/-} et *Hif1b*^{-/-} avec un rétrovirus surexprimant la protéine AID et les ai cultivées pendant 3 jours en présence de TGFb, IL4 et @CD40 pour induire la CI. J'ai constaté que la surexpression de AID dans les cellules *Hif1a*^{-/-} ou *Hif1b*^{-/-} est capable de restaurer la CI. Je conclus donc que le défaut de CI observé dans les cellules *Hif1a*^{-/-} et *Hif1b*^{-/-} est bien lié à une expression sous-optimale de l'AID.



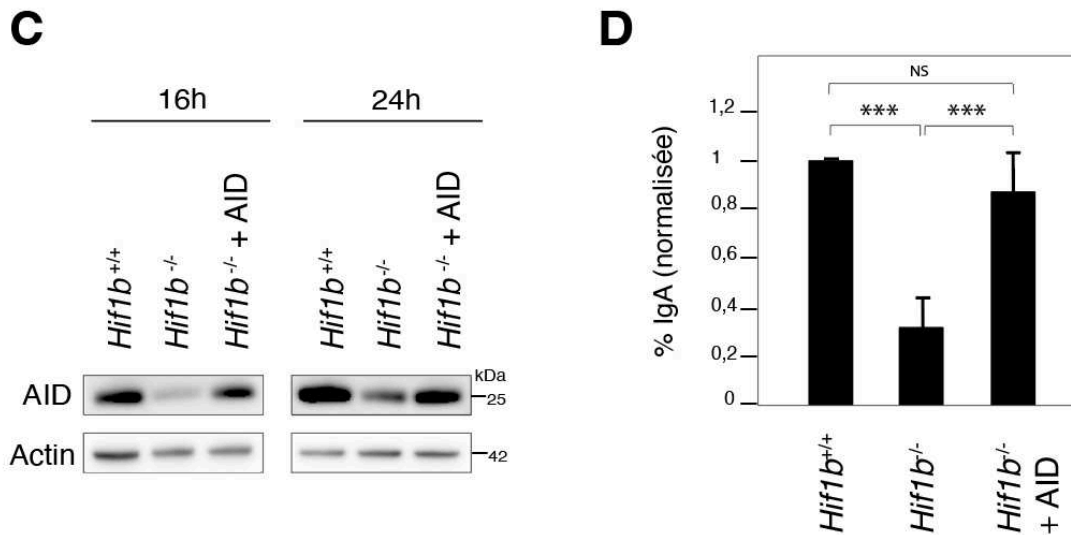


Figure 42: Le défaut de CI dans les cellules B déficientes en Hif1a ou en Hif1b est dû à l'expression sous-optimale d'AID.

Western blot pour la protéine AID et β -Actine dans (A) les cellules CH12 de type sauvage et *Hif1a*^{-/-} sur-exprimant AID ou non, et (C) les cellules CH12 de type sauvage, *Hif1b*^{-/-} et *Hif1b*^{-/-} sur-exprimant AID ou non. Les poids moléculaires (kDa) sont indiqués. Pourcentage de cellules exprimant IgA dans (B) les cellules *Hif1a*^{-/-} ou *Hif1b*^{-/-} (D), les cellules CH12 réexprimant AID (ou non), par rapport aux témoins de type sauvage. Les données proviennent de 5 expériences indépendantes. Valeur de p (test t de Student bilatéral; ***p<0,001, NS p>0,05); les valeurs de p sont *Hif1a*^{+/+} par rapport à *Hif1a*^{-/-} = 0,00018; *Hif1a*^{-/-} par rapport à *Hif1a*^{-/-} + AID = 0,00079; *Hif1a*^{+/+} par rapport à *Hif1a*^{-/-} + AID = 0,65873; *Hif1b*^{+/+} par rapport à *Hif1b*^{-/-} = 6,5x10⁻⁶; *Hif1b*^{-/-} par rapport à *Hif1b*^{-/-} + AID = 0,00066; *Hif1b*^{+/+} par rapport à *Hif1b*^{-/-} + AID = 0,14032.

Chapitre 3 - Discussion

1 - Identification des co-facteurs de AID

Initialement, j'ai entrepris un projet de recherche d'envergure en protéomique, visant à décrypter le rôle fonctionnel des protéines nucléaires associées à AID par le biais de la surexpression de la protéine de fusion AID exogène. L'analyse des protéines ainsi isolées a abouti à l'identification de plus de 300 candidats, jetant ainsi les bases d'une exploration approfondie de l'interaction entre AID et ses partenaires. Dans une seconde phase, notre approche s'est affinée dans le but d'identifier de nouveaux partenaires d'AID, en éliminant tout biais lié à la surexpression de protéine. Cette fois-ci, nous avons opté pour la biotinylation in vitro des partenaires d'AID. Les résultats de ces expérimentations, bien que non inclus dans le manuscrit initial, ont été significatifs, permettant l'identification de plus de 1000 protéines, élargissant ainsi considérablement notre compréhension des interactions moléculaires. Afin de cibler davantage les partenaires potentiels, nous avons croisé cette liste avec une troisième compilation de 700 protéines obtenue grâce à un criblage d'inactivation à l'échelle du génome, utilisant le système CRISPR/Cas9. Cette approche intégrative a permis d'affiner la sélection et d'identifier des candidats dont l'interaction avec AID peut jouer un rôle crucial. L'utilisation conjointe de ces technologies sophistiquées a enrichi la liste des protéines potentiellement interactives avec AID, avec une majorité désormais déjà identifiée. La caractérisation fonctionnelle des facteurs jusqu'alors inconnus devrait fournir des perspectives cruciales pour une meilleure compréhension de la régulation de l'action d'AID afin de générer une très grande diversité d'anticorps, rôle clé dans l'immunité humorale, mais aussi dans le cadre de son potentiel oncogénique important qui repose sur sa capacité à provoquer des dommages à l'ADN et provoquer des cancers. Cependant, certaines protéines potentiellement régulatrices d'AID échappent encore à l'identification par les méthodes susmentionnées. Leur expression sélective en des conditions physiologiques particulières défie les technologies de criblage utilisées jusqu'à présent. D'où l'idée novatrice d'explorer si des conditions particulières pourraient se manifester au sein des organes lymphoïdes secondaires, ouvrant ainsi de nouvelles voies pour la recherche dans ce domaine complexe.

Il est ainsi démontré que, au cours des réponses immunitaires, des zones d'hypoxie se développent au sein des centres germinaux, pouvant potentiellement influencer la réponse humorale. Les mécanismes physiologiques qui régissent ces processus dans le microenvironnement des GC restent largement inconnus. Il est donc plausible que la tension d'oxygène et la réponse physiologique à l'hypoxie jouent un rôle essentiel dans la régulation de la réponse humorale au sein des GC, HIF1 pourrait donc bien jouer un rôle dans la régulation de l'activation transcriptionnelle du gène AID au sein des cellules B, participant ainsi activement au processus de commutation isotypique au cours des réponses immunitaires. La question cruciale soulevée par cette conjecture est de savoir si *Hif1* est indispensable pour une commutation isotypique optimale au sein des cellules CH12. Cette interrogation a ouvert une voie intéressante pour explorer le lien entre la régulation de l'oxygène, la réponse à l'hypoxie, et les mécanismes cellulaires spécifiques qui peuvent coordonner la commutation isotypique dans le contexte des réponses immunitaires.

2 - Impact de HIF sur l'expression de AID et de la CI

J'ai montré que la CI est défectueuse dans les cellules B déficientes en *Hif1a* et *Hif1b*. J'observe que dans ces cellules il y a une expression retardée de l'ARNm et de la protéine AID qui est visible après 16h et qui dure jusqu'au pic d'expression de AID (24h). Ceci, en soi, explique le défaut de CI observé et me permet de conclure que la déficience du facteur de transcription HIF entraîne des niveaux sous-optimaux d'expression de AID et donc une CI défectueuse. Cette conclusion est étayée par le fait que la surexpression de AID contourne le défaut de CI observé dans un contexte déficient en *Hif1a* ou *Hif1b*.

Bien que le rôle fondamental de l'hypoxie et du facteur de transcription HIF-1 dans la biologie des cellules B semble clairement établi, il demeure difficile de préciser leur impact dans l'élaboration de la réponse humorale au sein de la réaction du centre germinatif in vivo. Cette complexité pourrait être attribuée en partie à l'effet pléiotropique de HIF-1 dans la régulation de divers aspects intrinsèques des cellules B en réponse à l'hypoxie, tels que le métabolisme, la survie cellulaire, la signalisation, ainsi que l'expression d'AID. (Zhang, 2022 ; Cho, 2016 ; Li, 2021).

Par ailleurs, l'hypoxie pourrait également exercer une influence sur la fonction des cellules auxiliaires folliculaires T et/ou des cellules dendritiques folliculaires, avec un impact potentiel sur la sélection des cellules B et la maturation d'affinité.

Il existe également des contradictions apparentes concernant l'impact de l'hypoxie sur les centres germinatifs et la diversification des anticorps par le biais de la CI et de l'hypermutation somatique (HMS). Des études antérieures (Rocco, 2019) ont démontré que la CI se produit au stade pré-GC en dehors des centres germinatifs. De plus, il a été observé que la région présentant une hypoxie plus marquée se trouve dans la zone claire plutôt que dans la zone sombre des centres germinatifs, où l'expression prédominante de AID est notée (Cho, 2016 ; Victoria, 2022).

Il est plausible que l'hypoxie déclenche la transactivation de l'ARNm d'AID dans la zone claire, atteignant son expression protéique maximale lorsque les cellules B parviennent à la zone sombre. De surcroît, étant donné que les organes lymphoïdes secondaires présentent généralement un environnement hypoxique par rapport à d'autres tissus (Burrows, 2017), cela pourrait inciter les cellules B des centres folliculaires et/ou germinatifs à renforcer une expression optimale d'AID.

En soutien au rôle du facteur de transcription HIF dans la régulation de l'expression d'AID, des études ont mis en évidence que l'inactivation de pVHL entraîne une accumulation significative de la protéine HIF-1 α , ainsi qu'une augmentation des niveaux de protéines des membres de la famille des cytidine désaminases APOBEC3 (A3B), parmi lesquels AID (Scholtès, 2021). On pourrait anticiper que l'inactivation de pVHL conduirait à une augmentation de l'activité transcriptionnelle dépendante de HIF. Ce renforcement pourrait induire une production accrue d'ARNm et de protéine AID, ainsi qu'une amélioration de l'efficacité de la CI et de l'HMS, en réaction à l'environnement hypoxique au sein du centre germinatif ou en réponse à la signalisation des cytokines.

Néanmoins, si un knock-out spécifique de pVHL dans les cellules B, a entraîné une stabilisation constitutive de HIF-1 α (Scholtès, 2021), il aurait étonnamment conduit à une altération de la maturation de l'affinité et à une CI défectueuse (Cho, 2016).

Ces résultats mettent en lumière la complexité inhérente à l'exploration du rôle de l'hypoxie et de HIF dans la diversification des anticorps *in vivo*.

Il est important de souligner que l'environnement hypoxique caractéristique de nombreux cancers entraîne une augmentation de HIF à l'intérieur de la cellule, ce qui pourrait également induire une augmentation de AID. Étant donné que la protéine AID possède un pouvoir oncogénique important, la compréhension de l'interaction entre l'expression de HIF et d'AID en environnement hypoxique dans les cellules semble donc pertinente dans un contexte de recherche en cancérologie.

Comme la voie de signalisation HIF-1 peut être activée (indépendamment de l'oxygène) par l'utilisation de cytokines, (Zhang, 2022 ; Li, 2021) notamment par TGF β , IL4 et @CD40, mes résultats soulignent la possibilité d'un rôle indépendant de l'oxygène du facteur de transcription HIF-1 dans l'expression de AID et donc dans la modulation de l'efficacité de la réponse humorale. Néanmoins, le rôle de HIF dans le HMS, qui nécessite la génération d'un modèle de souris knock-out conditionnel, reste à déterminer.

Enfin, étant donné que la diminution de l'expression d'AID dans les cellules *Hif1a*^{-/-} n'est pas entièrement abolie, contrairement aux cellules *Hif1b*^{-/-}, il est envisageable qu'il y ait une compensation fonctionnelle de *Hif1a* par *Hif2a*. Le rôle de *Hif2a* dans les cellules B n'est pas clairement établi et reste à déterminer (Zhang, 2022).

Le niveau de transcription et d'expression de AID subit des altérations significatives dans les cellules *Hif1a*^{-/-} à 16 heures, mais semble compensé à 24 heures, et surtout à 48 heures (Figure 43). En revanche, pour les cellules *Hif1b*^{-/-}, les deux niveaux demeurent très bas, sans manifestation apparente de compensation.

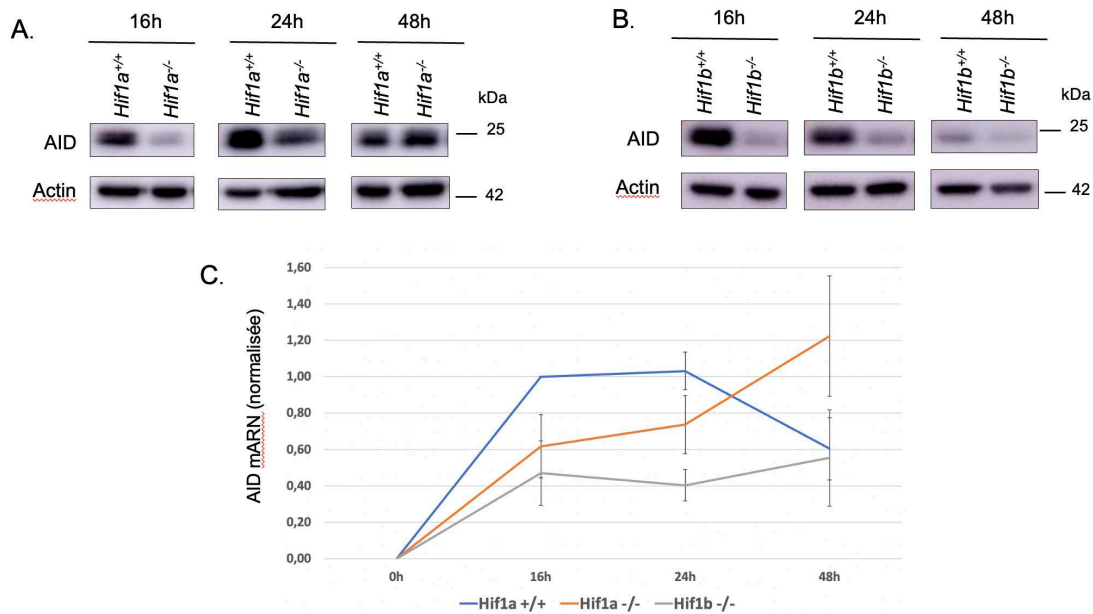


Figure 43: Compensation de l'expression de la protéine AID.

(A) Western blot pour la protéine AID et β -Actine dans (A) les cellules CH12 de type sauvage et *Hif1a*^{-/-} à 16, 24 et 48h, (B) dans les cellules *Hif1b*^{-/-}. Les poids moléculaires (kDa) sont indiqués. (C) Histogramme montrant les niveaux de transcription de l'Aicda dans les cellules sauvages, *Hif1a*^{-/-} et *Hif1b*^{-/-}, déterminés par RT-qPCR. Les valeurs moyennes des triplicats techniques de trois expériences indépendantes ont été normalisées par rapport aux transcrits HPRT (méthode des CT) et sont indiquées par rapport aux cellules WT, qui ont été normalisées à un.

En conclusion, mes recherches ont démontré que l'inactivation du facteur de transcription HIF dans les cellules B murines, indépendamment des effets pléiotropiques sur le métabolisme, la survie cellulaire ou la dépendance à l'oxygène, conduit à une CI défectueuse. Ce phénomène est attribuable à une induction sous-optimale de l'expression de AID, une découverte totalement innovante. Ces résultats revêtent des implications majeures pour l'efficacité de la réponse humorale dans les centres germinatifs, où les régions d'hypoxie et la stimulation par des cytokines pourraient potentialiser l'expression de AID, renforçant ainsi la diversification du récepteur des cellules B et contribuant à l'établissement d'une immunité hautement spécifique et adaptative. Plus largement, mes conclusions ouvrent de nouvelles perspectives de recherche sur le rôle de l'hypoxie et des voies immunitaires régulatrices induites par l'hypoxie, avec des implications potentielles pour le développement de stratégies de vaccination plus efficaces, et une meilleure compréhension du rôle d'AID comme oncogène.

Chapitre 4 - Perspectives

Comme évoqué précédemment, l'inactivation de pVHL conduit à une accumulation significative de la protéine HIF-1 α , ainsi qu'à une augmentation des niveaux de protéines des membres de la famille des cytidine désaminases APOBEC3 (A3B), parmi lesquels se trouve AID (Scholtes, 2021). Il est documenté que des conditions de culture hypoxique peuvent favoriser la commutation isotypique (CI) après la stimulation des lymphocytes B. Des expériences ont été menées avec des cellules stimulées (passant d'IgM à IgG1), placées en conditions normobariques (21% O₂) et hypoxiques (1% O₂) pendant 4 jours, démontrant une accélération de la cinétique de CI en milieu hypoxique. À l'inverse, une stimulation en environnement hyperoxique (60% O₂) entraîne une diminution de la production d'anticorps IgG1 (Abbott, 2016). Il serait donc intéressant d'explorer si l'expression de AID pourrait être augmentée en conditions d'hypoxie, et dans quelle mesure cela pourrait influencer la commutation isotypique. Cette ligne de recherche offre une perspective prometteuse pour approfondir notre compréhension des mécanismes moléculaires impliqués dans la réponse immunitaire, en mettant en lumière le potentiel impact de l'environnement hypoxique sur la régulation de AID et, par extension, sur les processus de commutation isotypique. Une telle exploration pourrait contribuer à éclairer les relations complexes entre les conditions de microenvironnement et la régulation génique dans le contexte des réponses immunitaires.


Il est important de noter que malgré le fait qu'un environnement hypoxique ne soit pas propice au développement adéquat des cellules, des recherches récentes indiquent qu'une exposition brève à une concentration élevée d'oxygène peut induire une réponse similaire à l'hypoxie, phénomène connu sous le nom de NOP (paradoxe de l'oxygène normobarique). Lors d'un stimulus hyperoxique bref (O₂ à 65% - pendant 2 heures par jour), la cellule saine est exposée à un stress oxydatif suffisant pour stimuler l'activité antioxydante, notamment à travers le système du glutathion, un puissant capteur des espèces réactives de l'oxygène (ROS) dans le corps humain. En maintenant un équilibre entre l'augmentation de ROS et l'augmentation des antioxydants dans la cellule, le stress oxydatif n'endommage pas la cellule pendant cette période d'hyperoxie courte. À leur retour à la normoxie, la chute de la pression

partielle d'oxygène est détectée par la cellule comme un stimulus hypoxique, déclenchant ainsi l'activation des mécanismes de réponse hypoxique, tels que l'expression du facteur HIF-1 α induit par l'hypoxie (Balestra, 2023). Il serait donc intéressant d'étudier si le NOP a un effet régulateur sur l'expression des protéines de la famille des HIF, sur AID et également sur la CI.

Dans un futur proche, il serait extrêmement intéressant de réaliser le knock-out de *Hif2a*, et surtout le double knock-out *Hif2a/Hif1a*, afin d'évaluer si une diminution continue de l'expression de AID serait observée. Cette approche permettrait d'approfondir notre compréhension des rôles spécifiques de HIF-2 α et HIF-1 α dans la régulation de AID, ce qui contribuerait également à notre compréhension de leur implication dans la commutation isotypique. De plus, étant donné que des mutations avec gain de fonction ont été démontrées au sein de la famille des HIFs, notamment pour *Hif2a*, il serait intéressant de déterminer si ces mutations pourraient éventuellement améliorer la diversification des immunoglobulines. Une telle investigation pourrait contribuer à élucider les mécanismes moléculaires sous-jacents à la modulation de la réponse immunitaire, ouvrant ainsi de nouvelles perspectives dans la compréhension des liens entre les mutations génétiques, les facteurs environnementaux, et la diversification des immunoglobulines au sein du système immunitaire.

Manuscrit

Short Communication

Optimal AID expression and efficient immunoglobulin class switch recombination are dependent on the hypoxia-inducible factorVincent Heyer^{1,2,3,4} and Bernardo Reina-San-Martin^{1,2,3,4} ¹ Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France² Institut National de la Santé et de la Recherche Médicale (INSERM), Illkirch, France³ Centre National de la Recherche Scientifique (CNRS), Illkirch, France⁴ Université de Strasbourg, Illkirch, France

During immune responses, B cells engaging a cognate antigen are recruited to GCs in secondary lymphoid organs where they will diversify their BCR to generate highly specific and adapted humoral responses. They do so, by inducing the expression of activation-induced cytidine deaminase (AID), which initiates somatic hypermutation (SHM) and class switch recombination (CSR). AID deaminates cytosines in ss DNA, generating U:G mismatches that are processed to induce ds DNA break intermediates during CSR that result in the expression of a different antibody isotype. Interestingly, hypoxia regions have been reported in GCs and suggesting that hypoxia could modulate the humoral response. Furthermore, hypoxia inducible transcription factor (HIF) can bind to the AID promoter and induce AID expression in a non-B-cell setting, suggesting that it might be involved in the transcriptional induction of AID in B cells, hence, regulating SHM and CSR. We, thus, hypothesized that HIF could regulate the efficiency of CSR. Here, we show that the inactivation of both the HIF-1 α and HIF-1 β subunits of the HIF transcription factor in murine CH12 B cells results in defective CSR and that this is due to the suboptimal induction of AID expression.

Keywords: AID · class switch recombination · B cell receptor · HIF-1 α · HIF-1 β 

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Introduction

During immune responses, the BCR repertoire is diversified in GC through somatic hypermutation (SHM) and class switch recombination (CSR) [1]. While SHM diversifies the variable region of Ig genes by introducing mutations and modifying antibody affinity for its cognate antigen, CSR provides novel antibody effector functions through a DNA recombination reaction (occurring at the Ig

heavy chain locus) that replaces the isotype expressed (from IgM to IgG, IgE, or IgA). Both of these processes are initiated in GC B cells by the inducible expression of activation-induced cytidine deaminase (AID). AID triggers these molecular events by deaminating cytosines to uracils in DNA, which are then processed by uracil DNA glycosylase and mismatch repair proteins to induce mutations or double stranded DNA breaks [1]. While AID-induced DNA lesions are essential to establish highly specific and adapted humoral responses, AID carries a significant oncogenic potential, which has been implicated at the origin of GC-derived B-cell lymphomas [2].

Correspondence: Dr. Bernardo Reina-San-Martin
e-mail: reinab@igbmc.fr

© 2023 The Authors. *European Journal of Immunology* published by Wiley-VCH GmbH

www.eji-journal.eu

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Given the strong oncogenic potential of AID, its expression and activity need to be tightly regulated. At the transcriptional level, AID expression is only induced in activated B cells and results from a tight balance between the activity of transcription factors (Batf, Hoxc4, etc.) and transcriptional repressors [3]. Interestingly, it has been previously shown that during immune responses, regions of hypoxia are observed within GCs and that the response to hypoxia modulates the humoral response [4, 5]. Nevertheless, little is known about the physiological mechanisms for regulating these processes within the GC microenvironment, and it is possible that oxygen tension and the physiological response to hypoxia may play a key role in regulating the humoral response within GCs [5–7].

The cellular response to hypoxia is mediated by the hypoxia inducible factor complex (HIF), formed by the HIF-1 α and HIF-1 β subunits. In normoxia, the HIF-1 α protein is unstable, it fixes the tumor suppressor factor von Hippel-Lindau (pVHL), an E3 ubiquitin ligase that ubiquitinates it and triggers its degradation through the proteasome. Hypoxia then triggers the stabilization of HIF-1 α , its association with HIF-1 β , and the nuclear translocation of the HIF complex, where it associates with other co-factors, such as p300/CBP, to bind to hypoxia response elements (HREs; RCGTG) and induce the transactivation of hypoxia-sensitive genes [5]. In addition, the HIF signaling pathway can also be activated in an oxygen-independent manner through the activation of the BCR, TLRs, or cytokine receptors, leading to the stabilization of the HIF-1 α protein in B cells [5].

It has been shown that HIF-1 α is capable of binding to the HREs in the *Aicda* promoter (located at -16 and -1620) inducing AID's mRNA production in a non-B-cell setting [8]. We, therefore, hypothesized that the HIF complex could be implicated in regulating the transcriptional activation of the *Aicda* gene in B cells. Here, we explore the role of HIF-1 α and HIF-1 β in the process of CSR in CH12 B cells.

Results and discussion

Hif1a is required for optimal CSR in CH12 cells

To determine whether HIF-1 α plays a role in CSR, we inactivated the *Hif1a* gene in CH12 cells (Supporting Information Fig. S1A), a murine B-cell line, that when cultured *in vitro* in the presence of TGF- β , IL4, and CD40 ligand is capable of undergoing CSR from IgM to IgA very efficiently [9]. The absence of HIF-1 α was verified by Western blot (Fig. 1A). To determine whether CSR is affected by inactivation of *Hif1a*, WT and *Hif1a*^{-/-} cells were induced to undergo CSR. After 72 h, the efficiency of CSR was evaluated by flow cytometry (Fig. 1B, C and Supporting Information Fig. S1D). While the parental cell line displayed 60% of IgA-expressing cells, this percentage was reduced by threefold in *Hif1a*^{-/-} CH12 cells (Fig. 1B,C). As CSR is dependent on proliferation, we induced CSR in CFSE-labeled cells and found that all genotypes diluted the CFSE dye at the same extent (Supporting Information Fig. S1C). To confirm that the defective CSR phenotype observed is

due to the lack of *Hif1a*, we transduced *Hif1a*^{-/-} cells with a retrovirus expressing HIF-1 α . We find that re-expression of HIF-1 α is able to rescue significantly the percentage of IgA-expressing cells (Fig. 1B,C). We conclude that *Hif1a* is required for optimal CSR in CH12 cells, independently of proliferation defects.

Induction of the AID mRNA is compromised in Hif1a^{-/-} cells

To determine whether AID expression is defective in the absence of *Hif1a*, we quantified the level of AID protein in mRNA by RT-qPCR (Fig. 1D) and Western blot (Fig. 1E). We found that induction of the AID mRNA is compromised in *Hif1a*^{-/-} cells (Fig. 1D) and that this results in delayed AID expression at the protein level (Fig. 1E) and a concomitant defect in CSR (Fig. 1B,C). Interestingly, we can observe that the re-expression of HIF-1 α can rescue the levels of AID mRNA (Fig. 1D) and protein (Fig. 1E).

Hif1b is required for optimal AID expression and CSR

The nuclear translocation of HIF-1 α is dependent on HIF-1 β . Thus, in the absence of HIF-1 β , we should observe a similar phenotype. To test this hypothesis, we inactivated the *Hif1b* gene in CH12 cells (Supporting Information Fig. S1B). Absence of HIF-1 β was verified by Western blot (Fig. 2A). To determine whether CSR is affected by inactivation of *Hif1b*, WT and *Hif1b*^{-/-} cells were induced to undergo CSR (Fig. 2B, C and Supporting Information Fig. S1D). While the parental cell line displayed 55% of IgA-expressing cells, this percentage was significantly reduced in *Hif1b*^{-/-} cells (Fig. 2B and C). As *Hif1a*^{-/-} cells, *Hif1b*^{-/-} cells proliferated at the same level as WT cells (Supporting Information Fig. S1C). To confirm that the defective CSR phenotype observed is due to the lack of HIF-1 β , we transduced *Hif1b*^{-/-} cells with a retrovirus expressing HIF-1 β . We find that re-expression of HIF-1 β is able to rescue the CSR defect (Fig. 2B, C). Phenocopying *Hif1a*-deficiency, *Hif1b*^{-/-} cells displayed delayed AID mRNA expression and protein levels (Fig. 2D, E). We conclude that *Hif1b* is required for optimal AID expression and CSR in CH12 cells, independently of proliferation defects.

The CSR defect in Hif1a- and Hif1b-deficient B cells is due to defective AID expression

If HIF-1 α and HIF-1 β are involved in mediating optimal levels of AID mRNA production, then the CSR defect observed in *Hif1a*^{-/-} and *Hif1b*^{-/-} cells should be rescued by overexpressing AID. To test this hypothesis, we transduced *Hif1a*^{-/-} and *Hif1b*^{-/-} cells with a retrovirus expressing AID (Fig. 3A,C) and cultured them for 3 days in the presence of TGF- β , IL4, and CD40 ligand to induce CSR. We find that overexpressing AID in *Hif1a*^{-/-} or *Hif1b*^{-/-} cells is able to rescue CSR (Fig. 3B,D). We conclude that the CSR

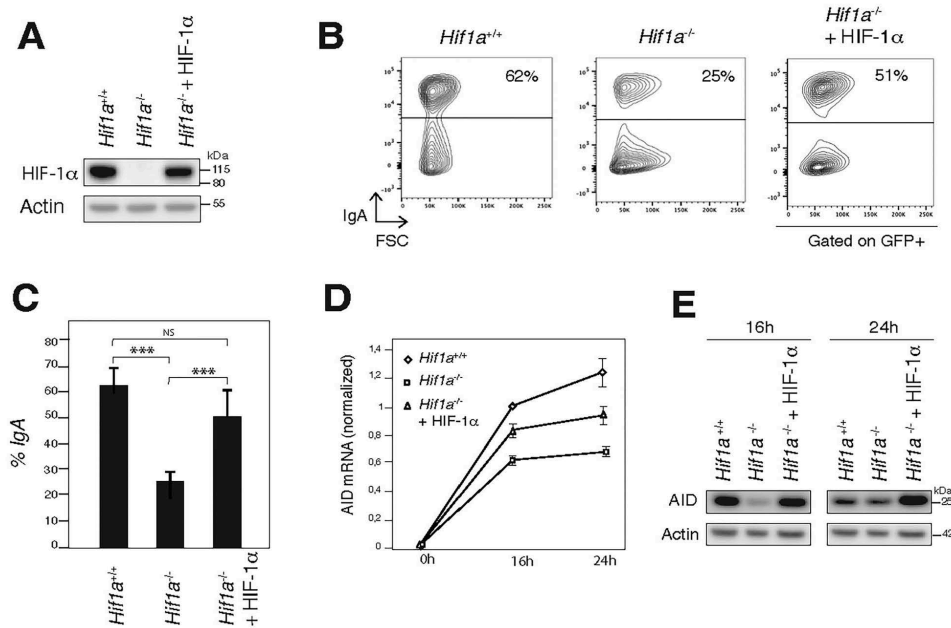


Figure 1. The hypoxia-inducible factor (HIF-1 α) is required for optimal AID expression and CSR. (A) Western blot for HIF-1 α , and β -Actin in WT, *Hif1 α ^{-/-}*, and *Hif1 α ^{-/-}* CH12 cells re-expressing HIF-1 α . Molecular weights (kDa) are indicated. (B) Flow cytometry analysis of IgA expression in WT, *Hif1 α ^{-/-}*, and *Hif1 α ^{-/-}* CH12 cells re-expressing HIF-1 α . Representative plots are shown. (C) Average percentage of IgA⁺ cells from five independent experiments in WT, *Hif1 α ^{-/-}*, and *Hif1 α ^{-/-}* CH12 cells re-expressing HIF-1 α . *p*-Value (two-tailed Student's *t*-test; ****p* < 0.001, NS *p* > 0.05); *p*-values are *Hif1 α ^{+/+}* versus *Hif1 α ^{-/-}* = 5.05×10^{-6} , *p*-values are *Hif1 α ^{-/-}* versus *Hif1 α ^{-/-}* + HIF-1 α = 0.00066, *p*-values are *Hif1 α ^{+/+}* versus *Hif1 α ^{-/-}* + HIF-1 α = 0.05264. (D) Histogram showing the *Aicda* transcript levels in WT, *Hif1 α ^{-/-}* and *Hif1 α ^{-/-}* CH12 cells re-expressing HIF-1 α as determined by RT-qPCR. The mean values of technical triplicates from three independent experiments were normalized to HPRT transcripts ($\Delta\Delta$ CT method) and are shown relative to WT cells, which were set to one. (E) Western blot analysis for AID and β -Actin in WT, *Hif1 α ^{-/-}*, and *Hif1 α ^{-/-}* CH12 cells re-expressing HIF-1 α . Molecular weights (kDa) are indicated. Uncropped Western blot images for Fig. 1A and E are shown in Supporting Information Fig. S2.

defect observed in *Hif1 α ^{-/-}* and *Hif1 β ^{-/-}* cells is related to sub-optimal expression of AID.

We have shown that CSR is defective in *Hif1 α* - and *Hif1 β* -deficient B cells. We observe that in these cells there is a delayed expression of the AID mRNA and protein that is visible after 16 h (Fig. 1D,E and 2D,E) and that lasts through the peak of AID expression (24 h). This, per se, explains the CSR defect observed (Fig. 1B,C and 2B,C) and allows us to conclude that deficiency in the HIF transcription factor results in suboptimal levels of AID expression and, hence, defective CSR. This conclusion is supported by the fact that overexpression of AID bypasses the CSR defect observed in a *Hif1 α* - or *Hif1 β* -deficient background (Fig. 3B,D).

While it is clearly established that hypoxia and the HIF1 transcription factor play a fundamental role in B-cell biology [5–7], it is difficult to pinpoint their precise role in shaping the humoral response at the GC reaction in vivo. This might be due, in part, to the pleiotropic effect of HIF1 in controlling intrinsic B-cell responses to hypoxia such as metabolism, cell survival, signaling, or AID expression [5, 7, 10]. In addition, hypoxia can also

influence the function of T-follicular helper cells and/or follicular DCs, thus, having a potential impact on B-cell selection and affinity maturation. Furthermore, there are some apparent contradictions concerning hypoxia in GCs and antibody diversification through CSR and SHM. For example, it has been shown that CSR occurs at the pre-GC stage outside of GCs [11] and that the region in which hypoxia is more pronounced is in the light zone [6] and not in the dark zone, where AID expression is predominant [12]. It is possible that hypoxia triggers the transactivation of the *Aicda* mRNA in the light zone and that its expression, at the protein level, is maximal once B cells reach the dark zone. Furthermore, as the secondary lymphoid organs are in general in a hypoxic environment relative to other tissues [7], this per se may poise follicular and/or GC B cells to enforce optimal AID expression.

Supporting the role of the HIF transcription factor in controlling AID expression, it has been shown that the inactivation of pVHL leads to a strong accumulation of the HIF-1 α protein, and consequently to increased protein levels of the members of the APOBEC3 cytidine deaminase (A3B) family, which includes AID

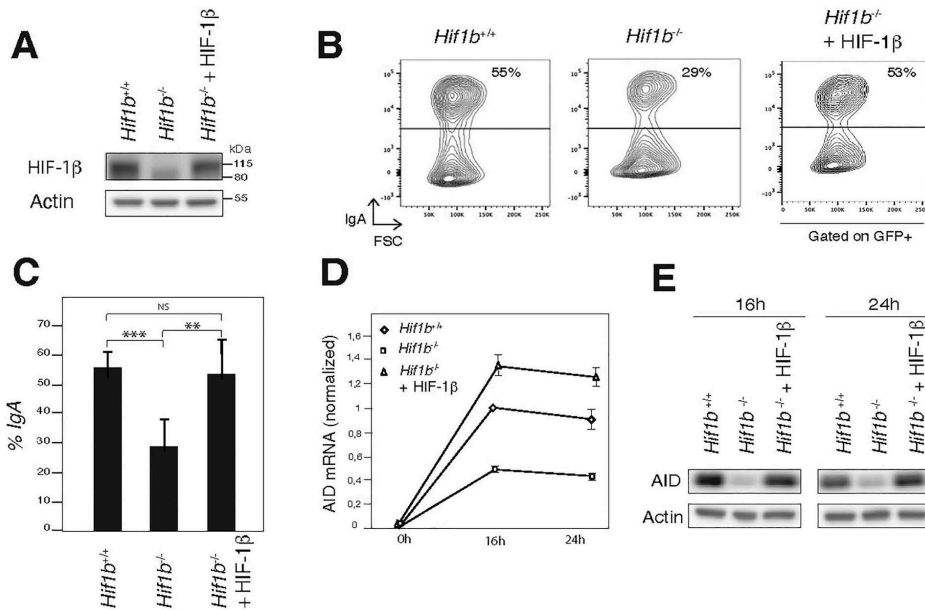


Figure 2. The hypoxia-inducible factor (HIF-1 β) is required for optimal AID expression and CSR. (A) Western blot for HIF-1 β and β -Actin in WT, *Hif1b*^{-/-}, and *Hif1b*^{-/-} CH12 cells re-expressing HIF-1 β . Molecular weights (kDa) are indicated. (B) Flow cytometry analysis of IgA expression in WT, *Hif1b*^{-/-}, and *Hif1b*^{-/-} CH12 cells re-expressing HIF-1 β . Representative plots are shown. (C) Average percentage of IgA⁺ cells from five independent experiments in WT, *Hif1b*^{-/-}, and *Hif1b*^{-/-} CH12 cells re-expressing HIF-1 β . p-Value (two-tailed Student's t-test; ****p* < 0.001, ***p* < 0.01, NS *p* > 0.05); p-values are *Hif1a*^{+/+} versus *Hif1a*^{-/-} = 0.0007, p-values are *Hif1a*^{-/-} versus *Hif1a*^{-/-} + HIF-1 α = 0.0079, p-values are *Hif1a*^{+/+} versus *Hif1a*^{-/-} + HIF-1 α = 0.75732. (D) Histogram showing the Aicda transcript levels in WT, *Hif1b*^{-/-}, and *Hif1b*^{-/-} CH12 cells re-expressing HIF-1 β as determined by RT-qPCR. The mean values of technical triplicates from three independent experiments were normalized to HPRT transcripts ($\Delta\Delta$ CT method) and are shown relative to WT cells, which were set to one. (E) Western blot analysis for AID and β -Actin in WT, *Hif1b*^{-/-}, and *Hif1b*^{-/-} CH12 cells re-expressing HIF-1 β . Molecular weights (kDa) are indicated. Uncropped Western blot images for Fig. 2A and E are shown in Supporting Information Fig. S2.

[13]. Thus, we would have expected that inactivation of pVHL would lead to enhanced HIF-dependent transcriptional activity [13], resulting in increased production of the AID mRNA and protein and more efficient CSR and SHM in a hypoxic environment in the GC or in response to cytokine signaling. Nevertheless, a B-cell-specific KO of pVHL, which resulted in constitutive HIF-1 α stabilization [6, 13], surprisingly led to impaired affinity maturation and defective CSR [6]. These results highlight the underlying difficulty of exploring the role of hypoxia and HIF in antibody diversification in vivo.

As the HIF signaling pathway can be activated (independently of oxygen) through cytokine signaling [5, 10], notably through TGF- β , IL4, and CD40, our results highlight the possibility of an oxygen-independent role of the HIF transcription factor in AID's expression and, hence, modulating the efficiency of the humoral response. Nevertheless, the role of HIF in SHM, which requires the generation of a KO mouse model, remains to be determined. In addition, as the efficiency of CSR in the absence of *Hif1a*, is not completely abolished, it is possible that there is a functional compensation by *Hif2a*, whose role in B cells is unclear [5].

While it has been shown that hypoxia leads to contrasting effects on CSR [4, 6, 14], here we demonstrate that inactivation of the HIF transcription factor in murine CH12 B cells (independently of pleiotropic effects on metabolism, cell survival, or oxygen dependence) results in defective CSR and that this is due to the suboptimal induction of AID expression. This has implications for the effectiveness of the humoral response within the GCs, as regions of hypoxia and cytokine stimulation might potentiate AID expression to enforce the diversification of the BCR and to accomplish the establishment of highly specific and adaptive immunity.

Materials and methods

Cell culture

CH12 cells were cultured in RPMI 1640 (Gibco) supplemented with 10 mM HEPES, 10% heat-inactivated Fetal Calf Serum (Bod-inco BV), 1 mM Na pyruvate, 100 U/mL penicillin/streptomycin,

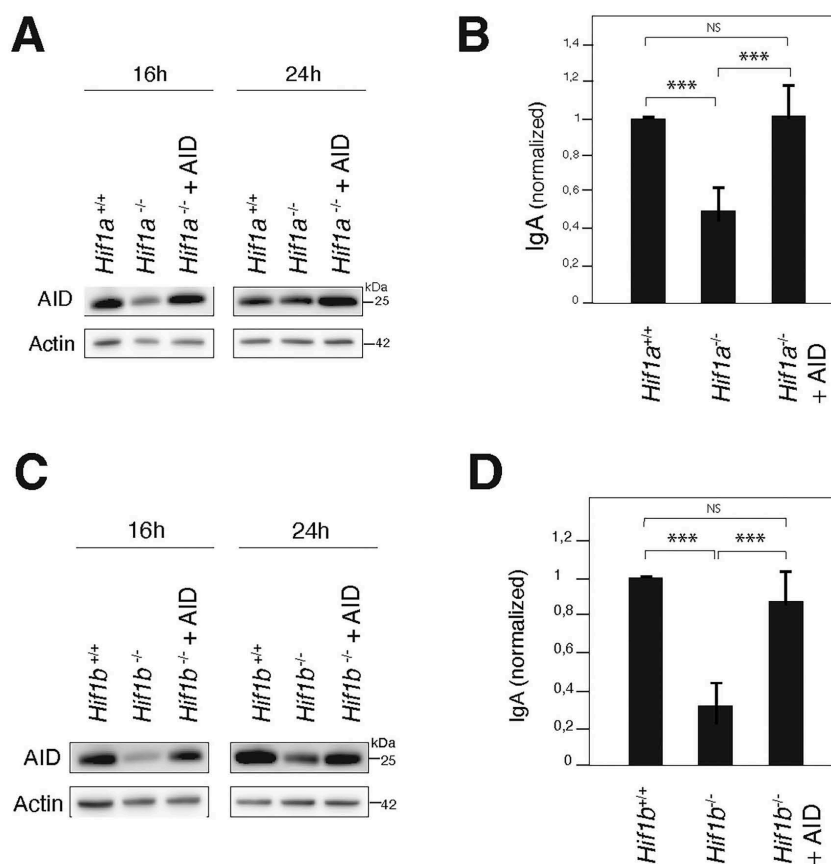


Figure 3. Overexpression of AID rescues the CSR defect in Hif1a- and Hif1b-deficient B cells. Western blot for AID and β -Actin in (A) WT and Hif1a^{-/-} CH12 cells overexpressing AID or not and (C) WT, Hif1b^{-/-}, and Hif1b^{-/-} CH12 cells overexpressing AID or not. Molecular weights (kDa) are indicated. Percentage of IgA-expressing cells in (B) Hif1a^{-/-} or Hif1b^{-/-} (D) CH12 cells re-expressing AID (or not), relative to WT controls. Data are from five independent experiments. *p*-Value (two-tailed Student's *t*-test; ****p* < 0.001, NS *p* > 0.05); *p*-values are Hif1a^{+/+} versus Hif1a^{-/-} = 0.00018; Hif1a^{-/-} versus Hif1a^{-/-} + AID = 0.00079; Hif1a^{+/+} versus Hif1a^{-/-} + AID = 0.65873; Hif1b^{+/+} versus Hif1b^{-/-} = 6.5×10^{-6} ; Hif1b^{-/-} versus Hif1b^{-/-} + AID = 0.00066; Hif1b^{+/+} versus Hif1b^{-/-} + AID = 0.14032. Uncropped Western blot images for Fig. 3A and C are shown in Supporting Information Fig. S2.

and 50 μ M β -mercaptoethanol. HIF-1 α protein was induced with 1% O₂, 35 μ g/mL COCl₂, or 5 mM DMOG, for 6 h. Bosc23 cells were cultured in DMEM (Gibco) supplemented with glucose (4.5g/L), 10% of heat-inactivated fetal calf serum, 1 mM Na pyruvate, 100 U/mL penicillin/streptomycin.

CRISPR/Cas9

CH12 cells were transfected with a plasmid expressing Cas9 high-fidelity (Cas9-HF1) fused to EGFP and co-expressing

2 gRNAs flanking exon 6 or exon 2 of the *hif1a* or *hif1b* genes, respectively (Hif1a-1: GCATGCCCTGATATACATGG, Hif1a-2: CTCATAGCTAAGATTACAGG, Hif1b-1: AATCCTCA-CACCCACATGG, Hif1b-2: CTCAGTGGTTAAGAGCACTG). After 24 h of transfection, single CH12 cells expressing Cas9-HF1 (EGFP+) were sorted (FACS Aria Fusion, BD) and cultured for 2 weeks. Clones were genotyped by PCR using specific primers (HIF1a-F: ATGACCAGATTAGTGAAGAAGGAAA-GATAG, HIF1a-R: TGTGTTTTAGTGTTTAGTTGGTTG, HIF1b-F: GAGTCAATTTTCATGTGGCCTGGATAGTG, HIF1b-R: CTGTGAGT-GAGGTCGGATGATAAGTAATAGGTGTG) and sequencing.

Western blot analysis

Whole-cell extracts were prepared using standard methods. Proteins were separated by SDS-PAGE, transferred to a PVDF membrane (Millipore) and blotted using anti HIF-1 α (1/1000; Novus NB100-123), anti-HIF-1 β (1/1000; Novus NB100-110), anti-AID (1/10.000; IGBMC), and anti- β -Actin (1/20.000; Sigma-Aldrich) antibodies.

CSR Assays

CH12 cells were cultured for 72 h in vitro with IL4 (5 μ g/ μ L; Peprotech), TGF- β (1 μ g/ μ L; R&D Systems Europe), and anti-CD40 antibody (200 ng/mL; eBioscience) for 72 h. CSR to IgA was assessed by staining with an anti-IgA-PE antibody (1/500; Southern Biotech). Before analysis, DAPI was added to exclude dead cells. Cells were analyzed by flow cytometry (Celesta BD) using the FlowJo software and following the flow cytometry guidelines. Gating strategy is shown in Supporting Information Fig. S1D.

Retroviral transductions

Cells were transduced with retroviral supernatants obtained by transfecting Bosc23 cells with a retrovirus expressing mHIF-1 α , mHIF-1 β , or mAID and selected with puromycin (1 μ g/mL). After 7 days, cells were submitted to CSR assays.

RT-qPCR

RNA extraction and cDNA synthesis were done using standard techniques. RT-qPCR was performed in triplicate (Roche LightCycler 480) using QuantiTect SYBR green PCR master mix (Qiagen) and specific primers (AID-F: GAAAGTCACGCTG-GAGACCG, AID-R: TCTCATGCCGTCCTTGG, HPRT-F: GTTG-GATACAGGCCAGACTTTGTTG, HPRT-R: GAITCAACTTGCCT-CATCTTAGGC).

Acknowledgements: We thank members of the Reina-San-Martin laboratory for discussions. C. Ebel and M. Phillips for assistance with cell sorting. B. Heller for cell culture. This work of the Interdisciplinary Thematic Institute IMCBio, as part of the ITI 2021–2028 program of the University of Strasbourg, CNRS and Inserm, was supported by IdEx Unistra (ANR-10-IDEX-0002), by SFRI-STRATUS project (ANR 20-SFRI-0012), and EUR IMCBio (ANR-17-EURE-0023) under the framework of the French Investments for the Future Program and by grants to B.R.S.M. from Fondation ARC, Institut National Contre le Cancer (INCA) and Fondation Recherche Médicale (FRM).

Conflict of interest: The authors declare no commercial or financial conflict of interest.

Author contributions: Conceptualization, V.H. and B.R.-S.-M.; methodology, V.H. and B.R.-S.-M.; investigation, V.H.; writing-original draft, V.H. and B.R.-S.-M.; writing-review & editing, V.H. and B.R.-S.-M.; supervision, B.R.-S.-M.; funding acquisition, B.R.-S.-M.

Data availability statement: The data that support the findings of this study are available from the corresponding author upon request.

Peer review: The peer review history for this article is available at <https://publons.com/publon/10.1002/eji.202350373>

References

- Feng, Y.Q., Seija, N., Di Noia, J. M. and Martin, A., AID in antibody diversification: there and back again (vol 41, pg 586, 2020). *Trends Immunol.* 2021. 42: 89–89.
- Nussenzweig, A. and M.C. Nussenzweig., Origin of chromosomal translocations in lymphoid cancer. *Cell* 2010. 141: 27–38.
- Tran, T.H., Nakata, M., Suzuki, K., Begum, N. A., Shinkura, R., Fagarasan, S., Honjo, T. et al., B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. *Nat. Immunol.* 2010. 11: 148–U67.
- Abbott, R.K., Thayer, M., Labuda, J., Silva, M., Philbrook, P., Cain, D. W., Kojima, H. et al., Germinal center hypoxia potentiates immunoglobulin class switch recombination. *J. Immunol.* 2016. 197: 4014–4020.
- Zhang, J.W., Wu, X., Ma, J., Long, K., Sun, J., Li, M. and Ge, L., Hypoxia and hypoxia-inducible factor signals regulate the development, metabolism, and function of B cells. *Front. Immunol.* 2022. 13. <https://doi.org/10.3389/fimmu.2022.967576>
- Cho, S.H., Raybuck, A. L., Stengel, K., Wei, M., Beck, T. C., Volanakis, E., Thomas, J. W. et al., Germinal centre hypoxia and regulation of antibody qualities by a hypoxia response system. *Nature* 2016. 537: 234–238.
- Burrows, N. and P.H. Maxwell, Hypoxia and B cells. *Exp. Cell. Res.* 2017. 356: 197–203.
- Liu, Y.Q., Mukhopadhyay, P., Pisano, M. M., Lu, X., Huang, L., Lu, Q. and Dean, D. C., Repression of Zeb1 and hypoxia cause sequential mesenchymal-to-epithelial transition and induction of Aicd, Oct4, and Dnmt1, leading to immortalization and multipotential reprogramming of fibroblasts in spheres. *Stem Cells* 2013. 31: 1350–1362.
- Rogier, M., Moritz, J., Robert, I., Lescale, C., Heyer, V., Abello, A., Martin, O. et al., Fmn72a enforces error-prone DNA repair during antibody diversification. *Nature* 2021. 600: 329–333.
- Li, L., Feng, C., Qin, J., Li, D., Liu, M., Han, S. and Zheng, B., Regulation of humoral immune response by HIF-1 α -dependent metabolic reprogramming of the germinal center reaction. *Cell. Immunol.* 2021. 367: 104409.
- Roco, J.A., Mesin, L., Binder, S. C., Nefzger, C., Gonzalez-Figueroa, P., Canete, P. F., Ellyard, J. et al., Class-switch recombination occurs infrequently in germinal centers. *Immunity* 2019. 51: 337–350 e7.
- Victoria, G.D. and M.C. Nussenzweig, Germinal centers. *Annu. Rev. Immunol.* 2022. 40: 413–442.

- 13 Scholtes, G.K., Sawyer, A. M., Vaca, C. C., Clerc, I., Roh, M., Song, C. and D'Aquila, R. T., The von Hippel-Lindau Cullin-RING E3 ubiquitin ligase regulates APOBEC3 cytidine deaminases. *Transl. Res.* 2021. **237**: 1–15.
- 14 Koers, J., Marsman, C., Steuten, J., Tol, S., Derksen, N. I. L., ten Brinke, A., van Ham, S. M. et al., Oxygen level is a critical regulator of human B cell differentiation and IgG class switch recombination. *Front. Immunol.* 2022. **13**: p. 1082154.

hypoxia inducible transcription factor · pVHL: tumor suppressor factor von Hippel-Lindau · HREs: hypoxia response elements

Full correspondence: Dr. Bernardo Reina-San-Martin, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France.
e-mail: reinab@igbmc.fr

Received: 12/1/2023

Revised: 31/3/2023

Accepted: 2/5/2023

Accepted article online: 4/5/2023

Abbreviations: **AID**: activation-induced cytidine deaminase · **SHM**: somatic hypermutation · **CSR**: class switch recombination · **HIF**:

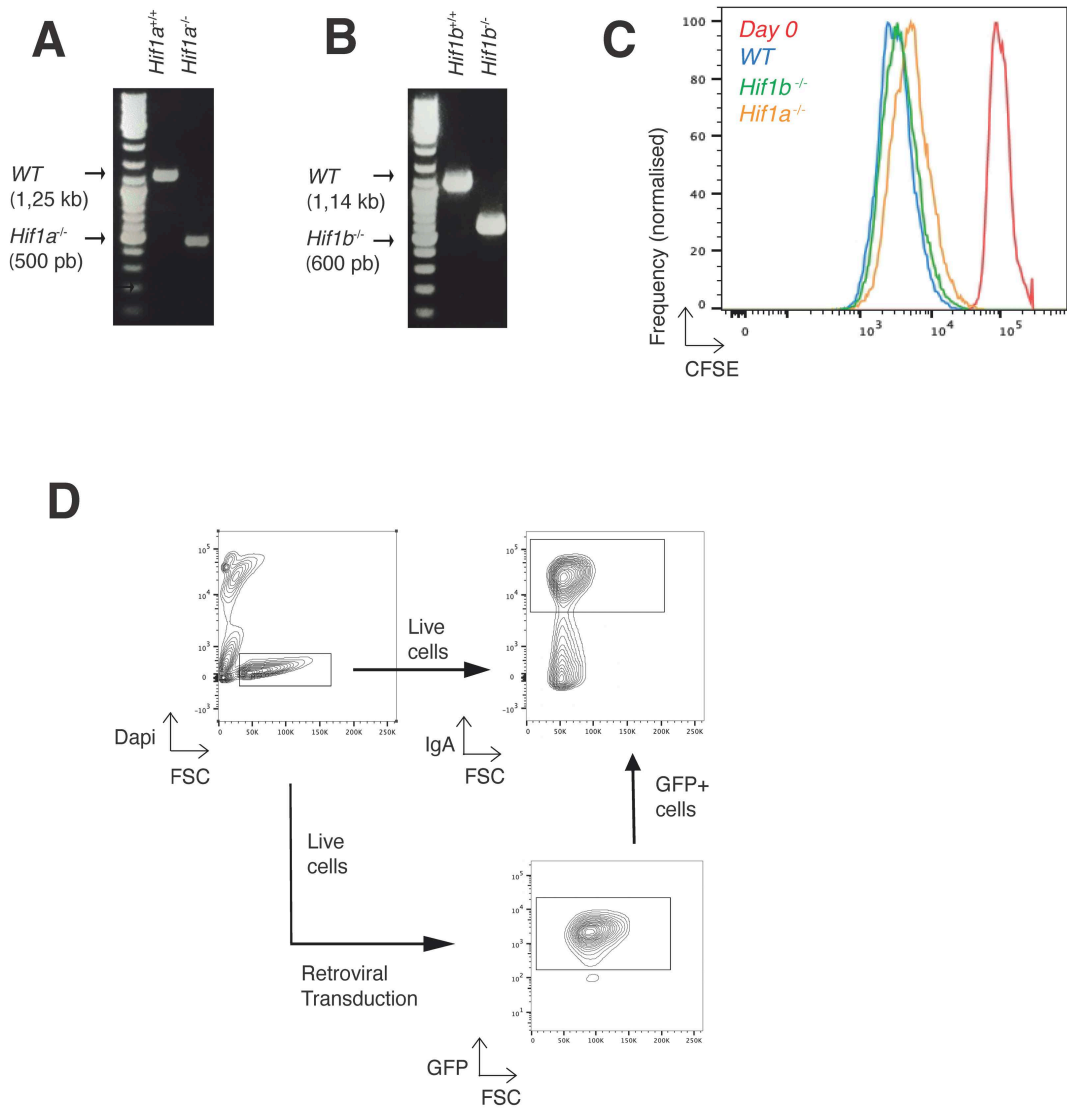
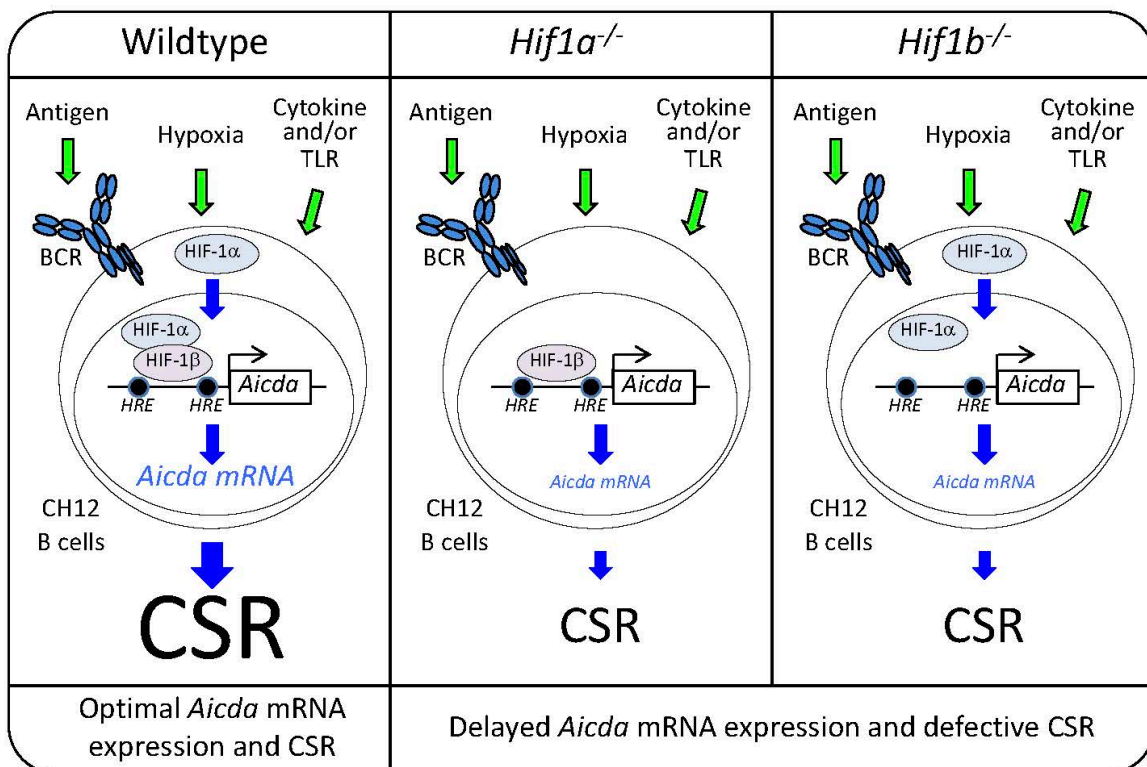


Figure S1. Genotyping by PCR of *Hif1a*^{-/-} (A) and *Hif1b*^{-/-} (B) CH12 cells clones. (C) CFSE dye dilution analysis by flow cytometry in wildtype, *Hif1a*^{-/-} and *Hif1b*^{-/-} CH12 cells. CFSE staining at day 0 is included for comparison. (D) Gating strategy related to all Figures.

Graphical Abstract



During immune responses, B cells respond to BCR, cytokine, TLR and/or hypoxia signaling to activate the HIF transcription factor and induce optimal levels of *Aicda* mRNA expression and CSR. In the absence of HIF-1a or HIF-1b, *Aicda* mRNA expression is delayed and results in defective CSR.

Matériel et Méthodes

Extraction des protéines cytoplasmiques et nucléaires:

Le culot de cellules CH12 (12 milliards de cellules) est repris dans le tampon de lyse cytoplasmique (10 mM Tris pH 7,9 ; 1,5 mM MgCl₂ ; 10 mM KCl ; 1mM DTT ; 0,5 mM PMSF). Après broyage des tissus (Wheaton dounce tissue grinder, B) et centrifugation, le surnageant contenant les protéines solubles est récupéré (S100). Le culot obtenu est lavé, puis repris dans le tampon de lyse nucléaire (20 mM Tris pH 7,9 ; 25% glycérol ; 1,5 mM MgCl₂ ; 500 mM KCl ; 1 mM EDTA ; 1 mM EGTA ; 1mM DTT ; 0,5 mM PMSF). Après broyage du culot, incubation 1 heure sous agitation et centrifugation, on récupère le surnageant contenant les protéines nucléaires (NE). S100 et NE sont dialysés contre le tampon BC100 (20 mM Tris pH 7,9 ; 20% glycérol ; 100 mM KCl ; 0,25 mM EDTA ; 0,125 mM EGTA ; 1mM DTT ; 0,5 mM PMSF ; 0,025% NP-40).

Culture cellulaire

Les cellules CH12 ont été cultivées dans du milieu de culture RPMI 1640 (Gibco) avec 10 mM HEPES, 10 % de sérum de veau foetal inactivé par la chaleur (Bod- inco BV), 1 mM de pyruvate de Na, 100 U/mL de pénicilline/streptomycine, et 50 µM de β-mercaptoéthanol. La protéine HIF-1α a été induite avec 1% d'O₂, 35 µg/mL de COCl₂, ou 5 mM de DMOG, pendant 6 h. Les cellules Bosc23 ont été cultivées dans du milieu de culture DMEM (Gibco) supplémenté en glucose (4,5g/L), 10% de sérum de veau foetal inactivé par la chaleur, 1 mM de pyruvate de Na, 100 U /mL de pénicilline/streptomycine.

CRISPR/Cas9

Les cellules CH12 ont été transfectées avec un plasmide exprimant Cas9 High-Fidelity (Cas9-HF1) fusionné à EGFP et coexprimant 2 gARN flanquant l'exon 6 ou l'exon 2 des gènes hif1a ou hif1b, respectivement :

Hif1a-1 : GCATGCCCTGATACATGG,

Hif1a-2 : CTCATAGCTAAGATTACAGG,

Hif1b-1 : AATCCTTCACACCCACATGG,

Hif1b-2 : CTCAGTGGTTAAGCACTG.

Après 24 h de transfection, les cellules CH12 exprimant Cas9-HF1 (EGFP+) ont été triées (FACS Aria Fusion, BD) et cultivées pendant 2 semaines. Les clones ont été génotypés par PCR en utilisant des amorces spécifiques :

HIF1a-F : ATGACCAGATTAGTGAAGAAGGAAAGATAG,

HIF1a-R : TGTGTTAGTGTTAGTTGGTTG,

HIF1b-F : GAGTCAATTCATGTGGCCTGGATAGTG,

HIF1b-R : CTGTGAGTGAGGTCGGATGATAAGTAATAGGTGTG)

Puis séquencés.

Analyse par Western blot

Les extraits de cellules totaux ont été préparés dans le tampon NuPage (LDS Sample Buffer de Invitrogen). Les protéines ont été séparées par SDS-PAGE (NuPage Bis Tris Mini Gel, Invitrogen), transférées sur une membrane PVDF (Millipore).

Les anticorps suivants ont été utilisés :

Anti-HIF-1 α (1/1000 ; Novus NB100-123),

Anti-HIF-1 β (1/1000 ; Novus NB100-110),

Anti-AID (1/10.000 ; IGBMC),

et anti- β -Actine (1/20.000 ; Sigma-Aldrich).

Stimulation des cellules à la CI

Les cellules CH12 ont été cultivées pendant 72 h in vitro avec de l'IL4 (5 μ g/ μ L ; Peprotech), du TGF- β (1 μ g/ μ L ; R&D Systems Europe) et un anticorps anti-CD40 (200 ng/mL ; eBioscience) pendant 72 h. La CI de IgM à l'IgA a été évaluée par marquage avec un anticorps anti-IgA-PE (1/500 ; Southern Biotech). Avant l'analyse, du colorant DAPI a été ajouté pour éliminer les cellules mortes. Les cellules marquées ont été analysées par cytométrie de flux (Fortessa BD) en utilisant le logiciel FlowJo.

Transductions rétrovirales

Les cellules ont été transduites avec des surnageants rétroviraux obtenus en transfectant les cellules Bosc23 avec un rétrovirus exprimant mHIF-1 α , mHIF-1 β ou mAID et sélectionné avec de la puromycine (1 μ g/mL). Après 7 jours de culture, les cellules ont été stimulées.

RT-qPCR

L'extraction de l'ARN a été réalisé par purification au Trizol/chloroforme, puis précipitation à l'isopropanol. La synthèse de l'ADNc ont été réalisées selon le protocole décrit dans le guide de la « SuperScript 4 First-Strand cDNA Synthesis Reaction de Lifetechnologies ».

La RT-qPCR a été réalisée en tripliquat (Roche LightCycler 480) en utilisant le master mix QuantiTect SYBR green PCR (Qiagen) et des amorces spécifiques suivantes :

AID-F : GAAAGTCACGCTGGAGACCG,

AID-R : TCTCATGCCGTCCTGG,

HPRT-F : GTTGGATACAGGCCAGACTTTGTTG,

HPRT-R : GATTCAACTTGCGCTCATCTTAGGC.

Références

- Abbott, R. K., Thayer, M., Labuda, J., Silva, M., Philbrook, P., Cain, D. W., Kojima, H., Hatfield, S., Sethumadhavan, S., & Ohta, A. (2016). Germinal center hypoxia potentiates immunoglobulin class switch recombination. *The Journal of Immunology*, *197*(10), 4014-4020.
- Burrows, N., & Maxwell, P. H. (2017). Hypoxia and B cells. *Experimental Cell Research*, *356*(2), 197-203.
- Chabi, S., Uzan, B., Naguibneva, I., Rucci, J., Fahy, L., Calvo, J., Arcangeli, M.-L., Mazurier, F., Pflumio, F., & Haddad, R. (2019). Hypoxia regulates lymphoid development of human hematopoietic progenitors. *Cell Reports*, *29*(8), 2307-2320. e2306.
- Chaudhuri, J., Basu, U., Zarrin, A., Yan, C., Franco, S., Perlot, T., Vuong, B., Wang, J., Phan, R. T., & Datta, A. (2007). Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Advances in immunology*, *94*, 157-214.
- Cho, S. H., Raybuck, A. L., Stengel, K., Wei, M., Beck, T. C., Volanakis, E., Thomas, J. W., Hiebert, S., Haase, V. H., & Boothby, M. R. (2016). Germinal centre hypoxia and regulation of antibody qualities by a hypoxia response system. *Nature*, *537*(7619), 234-238.
- Di Noia, J. M., & Neuberger, M. S. (2007). Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.*, *76*, 1-22.
- Jeevan-Raj, B. P., Robert, I., Heyer, V., Page, A., Wang, J. H., Cammas, F., Alt, F. W., Losson, R., & Reina-San-Martin, B. (2011). Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *Journal of Experimental Medicine*, *208*(8), 1649-1660.
- Li, L., Feng, C., Qin, J., Li, D., Liu, M., Han, S., & Zheng, B. (2021). Regulation of humoral immune response by HIF-1 α -dependent metabolic reprogramming of the germinal center reaction. *Cellular Immunology*, *367*, 104409.
- Liu, Y., Mukhopadhyay, P., Pisano, M. M., Lu, X., Huang, L., Lu, Q., & Dean, D. C. (2013). Repression of Zeb1 and hypoxia cause sequential mesenchymal-to-epithelial transition and induction of aid, Oct4, and Dnmt1, leading to immortalization and multipotential reprogramming of fibroblasts in spheres. *Stem Cells*, *31*(7), 1350-1362.
- Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., & Honjo, T. (2000). Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, *102*(5), 553-563.
- Muramatsu, M., Sankaranand, V., Anant, S., Sugai, M., Kinoshita, K., Davidson, N. O., & Honjo, T. (1999). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *Journal of Biological Chemistry*, *274*(26), 18470-18476.

- Nakamura, M., Kondo, S., Sugai, M., Nazarea, M., Imamura, S., & Honjo, T. (1996). High frequency class switching of an IgM⁺ B lymphoma clone CH12F3 to IgA⁺ cells. *International immunology*, 8(2), 193-201.
- Okazaki, I.-m., Hiai, H., Kakazu, N., Yamada, S., Muramatsu, M., Kinoshita, K., & Honjo, T. (2003). Constitutive expression of AID leads to tumorigenesis. *The Journal of experimental medicine*, 197(9), 1173-1181.
- Okazaki, I. m., Kotani, A., & Honjo, T. (2007). Role of AID in tumorigenesis. *Advances in immunology*, 94, 245-273.
- Petersen-Mahrt, S. K., Harris, R. S., & Neuberger, M. S. (2002). AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature*, 418(6893), 99-104.
- Rada, C., Jarvis, J. M., & Milstein, C. (2002). AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proceedings of the National Academy of Sciences*, 99(10), 7003-7008.
- Rajewsky, K. (1996). Clonal selection and learning in the antibody system. *Nature*, 381(6585), 751-758.
- Ramiro, A., San-Martin, B. R., McBride, K., Jankovic, M., Barreto, V., Nussenzweig, A., & Nussenzweig, M. C. (2007). The role of activation-induced deaminase in antibody diversification and chromosome translocations. *Advances in immunology*, 94, 75-107.
- Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Lagelouse, R., & Gennery, A. (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell*, 102(5), 565-575.
- Roco Alegre, J. E., Mesin, L., Binder, S. C., Nefzger, C. M., Gonzalez-Figueroa, P., Canete, P., Ellyard, J., Shen, Q., Robert, P. A., & Cappello, J. (2019). Class-Switch Recombination Occurs Infrequently in Germinal Centers.
- Rogier, M., Moritz, J., Robert, I., Lescale, C., Heyer, V., Abello, A., Martin, O., Capitani, K., Thomas, M., & Thomas-Claudepierre, A.-S. (2021). Fam72a enforces error-prone DNA repair during antibody diversification. *Nature*, 600(7888), 329-333.
- Scholtes, G. K., Sawyer, A. M., Vaca, C. C., Clerc, I., Roh, M., Song, C., & Richard, T. (2021). The von Hippel–Lindau Cullin-RING E3 ubiquitin ligase regulates APOBEC3 cytidine deaminases. *Translational Research*, 237, 1-15.
- Semenza, G. L. (2012). Hypoxia-inducible factors in physiology and medicine. *Cell*, 148(3), 399-408.
- Salvagno, M., Coppalini, G., Taccone, FS., Strapazzon, G., Mrakic-Sposta, S., Rocco, M., Khalife, M., Balestra, C. (2023). The Normobaric Oxygen Paradox. *Int J Mol Sci*, 24(1), 82.
- Thomas-Claudepierre, A.-S., Robert, I., Rocha, P. P., Raviram, R., Schiavo, E., Heyer, V., Bonneau, R., Luo, V. M., Reddy, J. K., & Borggreffe, T. (2016). Mediator facilitates transcriptional activation and dynamic long-range contacts at the IgH locus during class switch recombination. *Journal of Experimental Medicine*, 213(3), 303-312.

- Thomas-Claudepierre, A.-S., Schiavo, E., Heyer, V., Fournier, M., Page, A., Robert, I., & Reina-San-Martin, B. (2013). The cohesin complex regulates immunoglobulin class switch recombination. *Journal of Experimental Medicine*, 210(12), 2495-2502.
- Victoria, G. D., & Nussenzweig, M. C. (2022). Germinal centers. *Annual review of immunology*, 40, 413-442.
- Willmann, K. L., Milosevic, S., Pauklin, S., Schmitz, K.-M., Rangam, G., Simon, M. T., Maslen, S., Skehel, M., Robert, I., & Heyer, V. (2012). A role for the RNA pol II-associated PAF complex in AID-induced immune diversification. *Journal of Experimental Medicine*, 209(11), 2099-2111.
- Zhang, J., Wu, X., Ma, J., Long, K., Sun, J., Li, M., & Ge, L. (2022). Hypoxia and hypoxia-inducible factor signals regulate the development, metabolism, and function of B cells. *Frontiers in Immunology*, 13, 967576.

Annexes

Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination

Beena Patricia Jeevan-Raj,¹ Isabelle Robert,¹ Vincent Heyer,¹ Adeline Page,¹ Jing H. Wang,^{2,3,5} Florence Cammas,¹ Frederick W. Alt,^{2,3,4,6} Régine Losson,¹ and Bernardo Reina-San-Martin¹

¹Institut de Génétique et de Biologie Moléculaire et Cellulaire, Institut National de Santé et de Recherche Médicale Unité 964/ Centre National de Recherche Scientifique Unité Mixte de Recherche 7104, Université de Strasbourg, 67404 Illkirch, France

²Immune Disease Institute; ³Department of Medicine and ⁴Howard Hughes Medical Institute, Children's Hospital Boston;

⁵Department of Genetics; and ⁶Department of Genetics and Pediatrics; Harvard Medical School, Boston, MA 02115

Immunoglobulin class switch recombination (CSR) is initiated by double-stranded DNA breaks (DSBs) in switch regions triggered by activation-induced cytidine deaminase (AID). Although CSR correlates with epigenetic modifications at the IgH locus, the relationship between these modifications and AID remains unknown. In this study, we show that during CSR, AID forms a complex with KAP1 (KRAB domain-associated protein 1) and HP1 (heterochromatin protein 1) that is tethered to the donor switch region (S_μ) bearing H3K9me3 (trimethylated histone H3 at lysine 9) in vivo. Furthermore, in vivo disruption of this complex results in impaired AID recruitment to S_μ, inefficient DSB formation, and a concomitant defect in CSR but not in somatic hypermutation. We propose that KAP1 and HP1 tether AID to H3K9me3 residues at the donor switch region, thus providing a mechanism linking AID to epigenetic modifications during CSR.

CORRESPONDENCE

Bernardo Reina-San-Martin:
reinab@igbmc.fr

Abbreviations used: AID, activation-induced cytidine deaminase; ChIP, chromatin immunoprecipitation; CSR, class switch recombination; DDR, DNA damage response; DSB, double-stranded DNA break; HTS, high-throughput amplicon sequencing; IgH-FISH, IgH-specific fluorescence in situ hybridization; NHEJ, nonhomologous end joining; qRT-PCR, quantitative RT-PCR; SHM, somatic hypermutation.

During immune responses, B cells diversify their Ig genes by somatic hypermutation (SHM) and class switch recombination (CSR). SHM modifies antibody affinities by introducing mutations in the variable region of heavy (IgH) and light chain genes (Di Noia and Neuberger, 2007). CSR modulates antibody effector functions by replacing the antibody isotype expressed through a DNA recombination reaction that joins two switch regions (Chaudhuri et al., 2007). SHM and CSR are initiated by the deamination of cytosines to uracils in single-stranded DNA mediated by activation-induced cytidine deaminase (AID; Petersen-Mahrt et al., 2002; Chaudhuri et al., 2007; Di Noia and Neuberger, 2007). The resulting dU:dG mismatches are differentially processed to generate mutations in variable regions during SHM or double-stranded DNA breaks (DSBs) in switch regions during CSR (Chaudhuri et al., 2007; Di Noia and Neuberger, 2007). AID appears to find its targets by its binding to Spt5 and RNA

polymerase II (Pavri et al., 2010) and distinguishes between variable and switch regions by its association with proteins like RPA, PKA, 14-3-3, or by the formation of higher-order DNA structures in switch regions (Chaudhuri et al., 2007; Vuong et al., 2009; Xu et al., 2010; Yamane et al., 2011). AID itself also contributes to this choice, as N-terminal mutations in AID result in normal CSR but defective SHM (Shinkura et al., 2004; Wei et al., 2011) and conversely C-terminal truncations in AID result in normal SHM but defective CSR (Chaudhuri et al., 2007; Kracker et al., 2010). The latter phenotype is also observed in mice deficient for DNA damage response (DDR) components (Ramiro et al., 2007) and in a subset of hyper IgM patients with a CSR-specific defect (Kracker et al., 2010). It has been proposed that this domain associates with CSR-specific factors that could be required for targeting AID to switch regions (Ta et al., 2003) or for recombination downstream of DSB formation either by

J.H. Wang's present address is Integrated Dept. of Immunology, University of Colorado School of Medicine and National Jewish Health, Denver, CO 80206.

Dr. Losson died on 2 February 2010.

© 2011 Jeevan-Raj et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

mediating RNA editing (Doi et al., 2009; Nonaka et al., 2009) or by recruiting factors that promote efficient DNA repair (Barreto et al., 2003; Ta et al., 2003; Ito et al., 2004; Shinkura et al., 2004; Kracker et al., 2010). In addition, epigenetic modifications at the IgH locus, including histone H3 trimethylation at lysine 9 (H3K9me3), have been suggested to target the CSR machinery to switch regions (Wang et al., 2006, 2009; Chowdhury et al., 2008; Kuang et al., 2009). However, a direct causal link between these modifications and AID function in CSR has not been established.

RESULTS

AID associates with KAP1 (KRAB domain-associated protein 1)

To identify nuclear proteins associating with AID, we performed sequential immunoprecipitations followed by mass spectrometry identification using nuclear extracts prepared from CH12 B cells expressing tagged AID (AID^{Flag-HA}; Fig. S1). As negative controls, we used cell lines expressing the tags alone (Flag-HA) or an irrelevant tagged protein (EGFP^{Flag-HA}). Among the identified proteins specifically coprecipitating with AID but not with the negative controls, we found known AID partners, DNA-PK_{CS} (Wu et al., 2005) and CRM1 (Brar et al., 2004; Ito et al., 2004; McBride et al., 2004; Geisberger et al., 2009), proteins previously implicated in CSR, Ikaros (Sellars et al., 2009), MSH2 (Chaudhuri et al., 2007), and MSH6 (Chaudhuri et al., 2007), and proteins with no known function in SHM or CSR, like KAP1 (Fig. S1). KAP1 (also known as TRIM28, KRIP1, or Tif1 β) was selected for analysis because it has been previously implicated as an effector of the DDR (Ziv et al., 2006). KAP1 is also a transcriptional co-repressor (Cammass et al., 2000) that associates with members of the HP1 (heterochromatin protein 1) family to participate in chromatin packaging and heterochromatin formation (Nielsen et al., 1999). The association between AID and KAP1 was confirmed by reciprocal immunoprecipitations and Western blotting (Fig. 1 A) and was specific, as KAP1 did not coprecipitate with EGFP^{Flag-HA} (Fig. 1 B), despite relatively higher expression of EGFP^{Flag-HA} (Fig. S1). To determine whether AID associates with KAP1 through its C terminus, we performed immunoprecipitation and Western blot experiments using extracts prepared from CH12 cells expressing an AID C-terminal truncation mutant (AID^{Flag-HA} Δ 182–198). Although the expression level of AID^{Flag-HA} Δ 182–198 is significantly

lower than AID^{Flag-HA} or EGFP^{Flag-HA} (Fig. S1), we found that C-terminal deletion in AID did not disrupt its association with KAP1 (Fig. 1 C). To determine whether the association between AID and KAP1 can take place within chromatin, we performed a coimmunoprecipitation experiment on chromatin fractions prepared from CH12 B cells expressing tagged AID^{Flag-HA}. We found that chromatin-bound AID reciprocally coimmunoprecipitated with KAP1 (Fig. 1 D). We conclude that endogenous KAP1 associates with nuclear and chromatin-bound tagged AID through interactions that do not require its C-terminal domain.

KAP1 is required for efficient CSR but is dispensable for SHM

To determine whether KAP1 plays a role in CSR and SHM, we bred *Kap1* floxed mice (*Kap1*^{F/F}; Cammas et al., 2000) with *CD19*^{Cre/+} transgenic mice (Rickert et al., 1997) to inactivate *Kap1* in developing B cells. Despite efficient Cre-mediated deletion, no differences were observed in cellularity in the bone marrow or the spleen (not depicted), and B cell development appeared to be unaffected (Fig. S2). To determine whether CSR is affected by *Kap1* deficiency, we cultured in vitro CFSE-labeled splenic B cells isolated from *CD19*^{Cre/+}*Kap1*^{F/F} and control mice under conditions that induce CSR to various isotypes. We found that *Kap1* deficiency had no effect on proliferation (Fig. 2 A) or survival (not depicted). However, *Kap1* deficiency resulted in a 40–50% reduction in CSR to all isotypes tested (Fig. 2, A–C). Consistent with this, the level of postrecombination I μ -C_H transcripts, which appear only after recombination, was significantly reduced in *CD19*^{Cre/+}*Kap1*^{F/F} B cells (Fig. 2 C and Fig. S3). We conclude that *Kap1* deficiency results in a B cell-intrinsic defect in CSR that is independent of abnormalities in survival or proliferation.

To determine whether SHM is also affected by *Kap1* deficiency, we immunized *CD19*^{Cre/+}*Kap1*^{F/F} and control mice with NP-CGG, sorted germinal center B cells from the lymph nodes, and analyzed J_H4 intron sequences for mutations (Jolly et al., 1997). We did not find significant differences in mutation frequency (Fig. 2 D; *P* = 0.12), distribution, or pattern (Fig. S4) between sequences obtained from *CD19*^{Cre/+}*Kap1*^{F/F} (*n* = 83) and control (*n* = 91) animals. We conclude that KAP1 is involved in mediating CSR but appears to be dispensable for SHM.

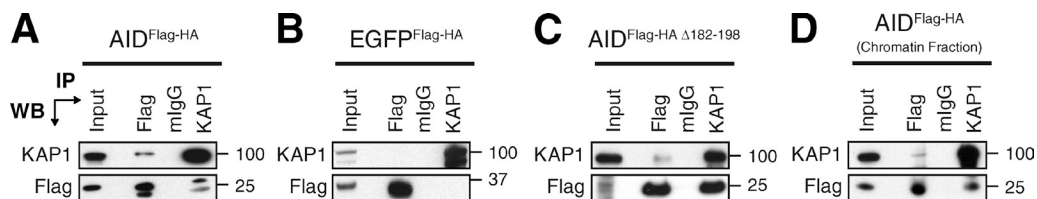


Figure 1. KAP1 associates with AID. (A–C) Nuclear extracts prepared from CH12 cells stably expressing AID^{Flag-HA} (A), EGFP^{Flag-HA} (B), or AID^{Flag-HA} Δ 182–198 (C) were immunoprecipitated and blotted with anti-KAP1 and anti-Flag antibodies. Because of the lower expression levels of AID^{Flag-HA} Δ 182–198 when compared with AID^{Flag-HA} or EGFP^{Flag-HA} (Fig. S1), loading for Western blot (WB) analysis was adjusted accordingly in C. (D) Chromatin fractions prepared from CH12 cells expressing AID^{Flag-HA} were immunoprecipitated and blotted with anti-KAP1 and anti-Flag antibodies. Input control represents 1% of material used for immunoprecipitation (IP). Molecular mass markers in kilodaltons are indicated. Data are representative of five independent experiments.

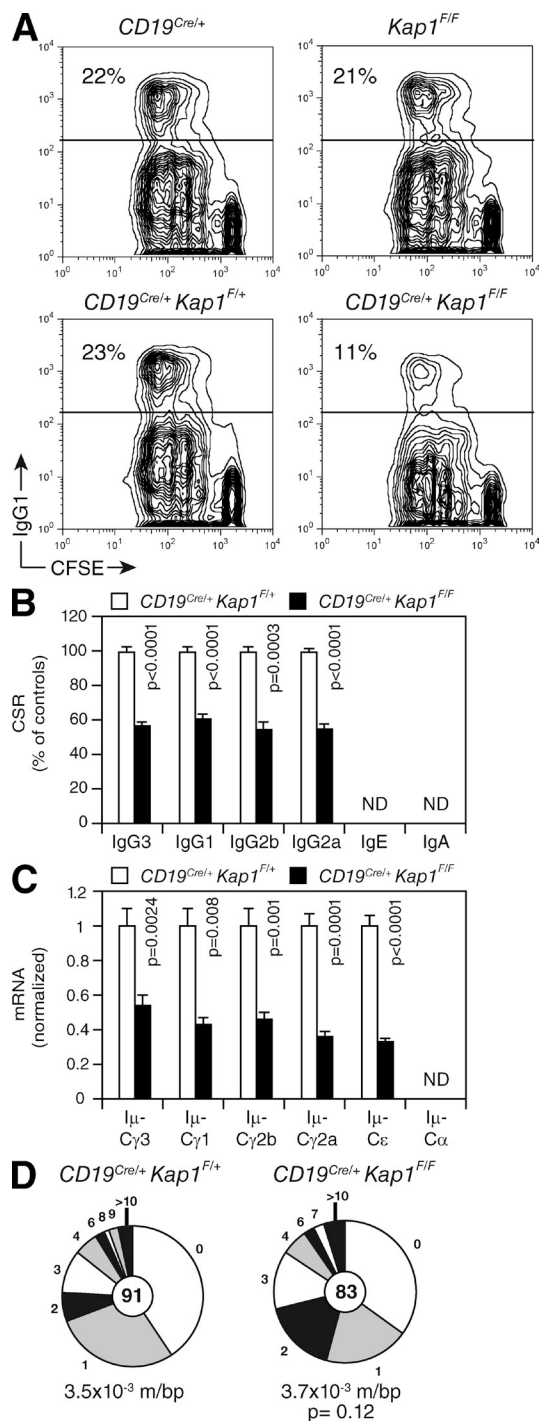


Figure 2. KAP1 is required for efficient CSR but is dispensable for SHM. (A) IgG1 surface expression and CFSE dilution by flow cytometry in $CD19^{Cre/+}$, $Kap1^{F/F}$, $CD19^{Cre/+}Kap1^{F/+}$, and $CD19^{Cre/+}Kap1^{F/F}$ B cells stimulated with LPS + IL-4 for 72 h. The percentage of switched cells is indicated in each plot. (B) Percentage (+SD) of CSR after 72 h in $CD19^{Cre/+}Kap1^{F/F}$ relative to $CD19^{Cre/+}Kap1^{F/+}$ for the different isotypes tested (CSR to IgA and IgE could not be determined). CSR in $CD19^{Cre/+}Kap1^{F/+}$ B cells was set to 100%. Data are from five independent experiments. (C) qRT-PCR for post-switch ($I\mu-C_H$) transcripts in $CD19^{Cre/+}Kap1^{F/F}$ relative to $CD19^{Cre/+}Kap1^{F/+}$ stimulated B cells (72 h). Expression is normalized to *Cd79b* expression and

The function of KAP1 in CSR is upstream of DSB resolution

CSR is dependent on transcription (Chaudhuri et al., 2007), and KAP1 is a transcriptional corepressor (Cammass et al., 2000). To determine whether switch region transcription is affected by *Kap1* deficiency, we measured by quantitative RT-PCR (qRT-PCR) the level of switch region germline transcripts in activated B cells. We found no significant reduction in the level of these transcripts in *Kap1*-deficient B cells (Fig. 3 A and Fig. S3). However, occasional increases in the levels of $\gamma 3$, $\gamma 1$, $\gamma 2b$, and $\gamma 2a$ transcripts were found (Fig. 3 A and Fig. S3). This is similar to what is observed in *Atm*^{-/-}, *H2ax*^{-/-}, or *53bp1*^{-/-} B cells in which CSR is reduced despite normal or increased levels of switch region transcripts (Reina-San-Martin et al., 2003, 2004; Ward et al., 2004). We conclude that switch regions continue to be efficiently transcribed in *Kap1*-deficient B cells and that switch regions are accessible for DNA deamination by AID. Therefore, the function of KAP1 is downstream of switch region transcription.

To determine whether *Kap1* deficiency affects AID expression, we measured the level of AID protein in activated B cells by Western blot. For comparison, we used *AID*-deficient (*AID*^{Cre/Cre}) B cells, which were obtained from mice bearing a targeted insertion of the CRE recombinase cDNA into *AID*'s exon 1 (Robbiani et al., 2008) and which displayed the same phenotype as *AID*^{-/-} mice (Muramatsu et al., 2000). Reduced levels in AID protein were found in *AID*^{Cre/+} B cells (Fig. 3 B), as has been reported for *AID*^{+/-} B cells (Serránchez et al., 2008). In contrast, *Kap1* deficiency had no deleterious effect on AID protein levels (Fig. 3 B). Consistent with this, retroviral overexpression of full-length AID in *Kap1*-deficient B cells did not rescue CSR, whereas it did in *AID*-deficient B cells (Fig. S5). As expected, overexpression of C-terminal truncation AID mutants had no effect on CSR in either *AID*- or in *Kap1*-deficient B cells (Fig. S5). We conclude that defective CSR in *Kap1*-deficient B cells is not caused by reduced levels of AID and that KAP1 functions downstream of AID expression.

KAP1 is phosphorylated at serine 824 (γ -KAP1) by the ATM kinase and accumulates in γ -H2AX-containing foci in response to DNA damage (Ziv et al., 2006). Furthermore, inactivation of DDR components (i.e., ATM and H2AX) results in defective CSR (Ramiro et al., 2007). To determine whether KAP1 is phosphorylated during CSR, we assayed for γ -KAP1 by Western blot. KAP1 phosphorylation was readily detected in control cells exposed to the DSB-inducing chemical neocarzinostatin (Fig. 3 C). Despite robust levels

is presented relative to the expression in $CD19^{Cre/+}Kap1^{F/+}$ B cells, set as 1. Data are representative of three experiments. Error bars indicate SD. (D) J_H4 intron sequences from germinal center B cells (B220⁺Fas⁺GL-7⁺) sorted from the lymph nodes of immunized $CD19^{Cre/+}Kap1^{F/F}$ and $CD19^{Cre/+}Kap1^{F/+}$ mice were analyzed for mutations. Segment sizes are proportional to the number of sequences bearing the indicated mutations. Mutation frequency (mutations/base pair [m/bp]) and number of sequences analyzed (in center) are indicated. Sequences were obtained from four independent experiments. (B–D) P-values were determined by the Student's *t* test.

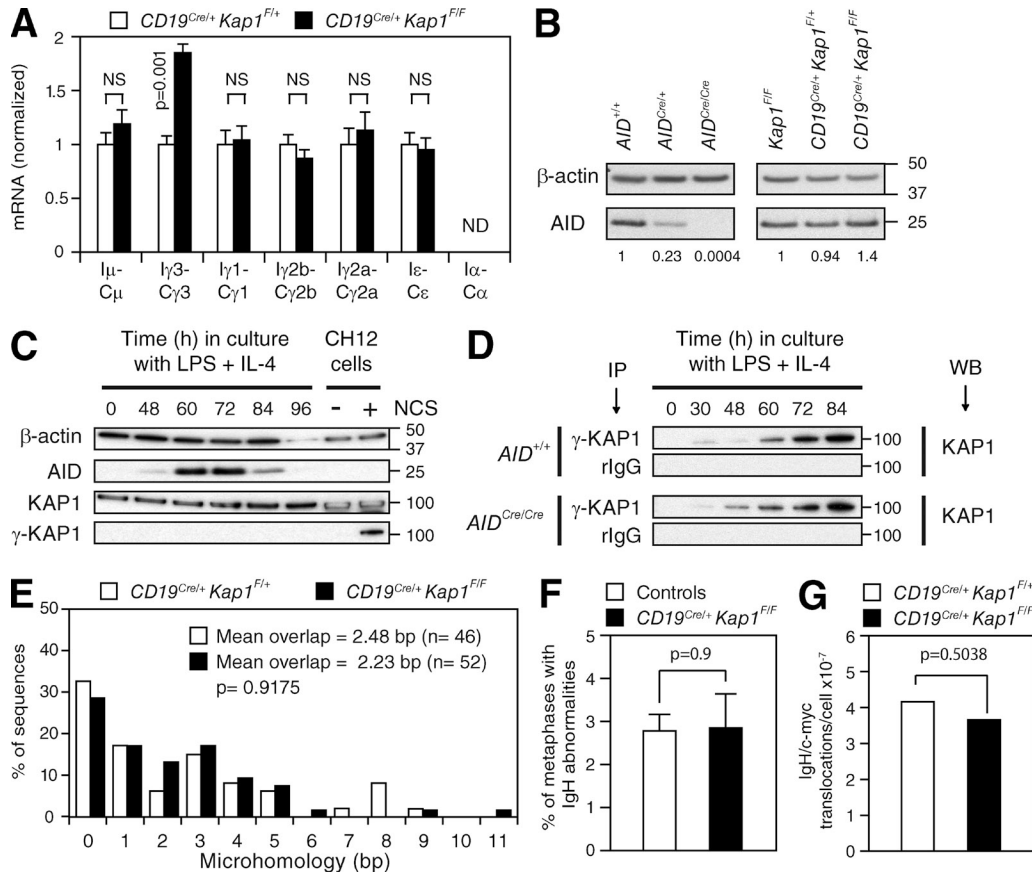


Figure 3. KAP1 functions downstream of switch region transcription and upstream of DSB formation. (A) qRT-PCR for germline (I_H-C_H) transcripts in $CD19^{Cre/+}Kap1^{F/F}$ relative to $CD19^{Cre/+}Kap1^{F/+}$ stimulated B cells (72 h). Expression is normalized to *Cd79b* and relative to $CD19^{Cre/+}Kap1^{F/+}$ B cells, set as 1. P-values were determined by the Student's *t* test. Data are from three independent experiments. Error bars indicate SD. (B) Western blot for AID and β -actin in $Kap1^{F/F}$, $CD19^{Cre/+}Kap1^{F/+}$, $CD19^{Cre/+}Kap1^{F/F}$, $AID^{+/+}$, $AID^{Cre/+}$, and $AID^{Cre/Cre}$ B cells. Numbers below the panel reflect the intensity of bands representing AID protein levels relative to controls after normalization to β -actin. Molecular mass markers in kilodaltons are indicated. Data are representative of three independent experiments. (C) KAP1 phosphorylation (γ -KAP1) by Western blot in wild-type B cells cultured with LPS + IL-4. As a positive control, CH12 cells treated with neocarzinostatin (NCS). (D) γ -KAP1 was immunoprecipitated from activated $AID^{+/+}$ and AID -deficient ($AID^{Cre/Cre}$) B cells and blotted for KAP1. Data are representative of two independent experiments. IP, immunoprecipitation; WB, Western blot. (E) Percentage of $S\mu/S\gamma3$ switch junction sequences with indicated nucleotide overlap from $CD19^{Cre/+}Kap1^{F/+}$ and $CD19^{Cre/+}Kap1^{F/F}$ LPS-stimulated B cells (72 h) from five independent experiments. Mean length of overlap in base pairs and the number of sequences analyzed (*n*) is indicated. (F) Quantification of IgH locus breaks as determined by IgH-FISH on metaphases prepared from control ($n = 780$) and $CD19^{Cre/+}Kap1^{F/F}$ ($n = 593$) B cells cultured for 3 d with anti-CD40 and IL-4. Mean (+SEM) is shown. Data are from two independent experiments (Table S1). (E and F) P-values were determined by the Mann-Whitney test. (G) Frequency of IgH/c-myc translocations in $CD19^{Cre/+}Kap1^{F/F}$ and $CD19^{Cre/+}Kap1^{F/+}$ B cells determined by long-range PCR and Southern blot in three independent experiments. The p-value was determined by the one-tailed exact Fisher's test.

of AID expression, γ -KAP1 was not detected in activated B cells (Fig. 3 C), indicating that KAP1 phosphorylation, if any, is below detection threshold. To determine whether a small fraction of KAP1 is phosphorylated during CSR, we immunoprecipitated γ -KAP1 and blotted for KAP1 (Fig. 3 D). Although γ -KAP1 was detected under these conditions, KAP1 phosphorylation was not dependent on AID expression (Fig. 3 D). We conclude that the role of KAP1 in CSR is independent of its phosphorylation status at serine 824.

CSR requires the generation of DSBs in switch regions and their resolution by nonhomologous end joining (NHEJ; Yan et al., 2007). To determine whether CSR junctions are affected by *Kap1* deficiency, we compared IgG3 CSR junctions from $CD19^{Cre/+}Kap1^{F/F}$ ($n = 52$) and control ($n = 46$) B cells.

Sequence analysis revealed no significant differences ($P = 0.9175$) in the amount of microhomology at the junctions between groups (Fig. 3 E). We conclude that CSR junctions are unaffected by *Kap1* deficiency and that DNA ends are repaired normally by NHEJ.

Inactivation of core NHEJ or DDR components results in global and IgH-specific genomic instability (Franco et al., 2006; Yan et al., 2007). To determine whether *Kap1* deficiency results in the accumulation of unresolved DNA breaks triggered by AID in switch regions, we performed two-color IgH-specific fluorescence in situ hybridization (IgH-FISH; Franco et al., 2006; Yan et al., 2007) on metaphase spreads prepared from $CD19^{Cre/+}Kap1^{F/F}$ ($n = 593$) and control ($n = 780$) activated B cells. Contrary to *Atm*^{-/-} or *Xrcc4*^{-/-} B cells

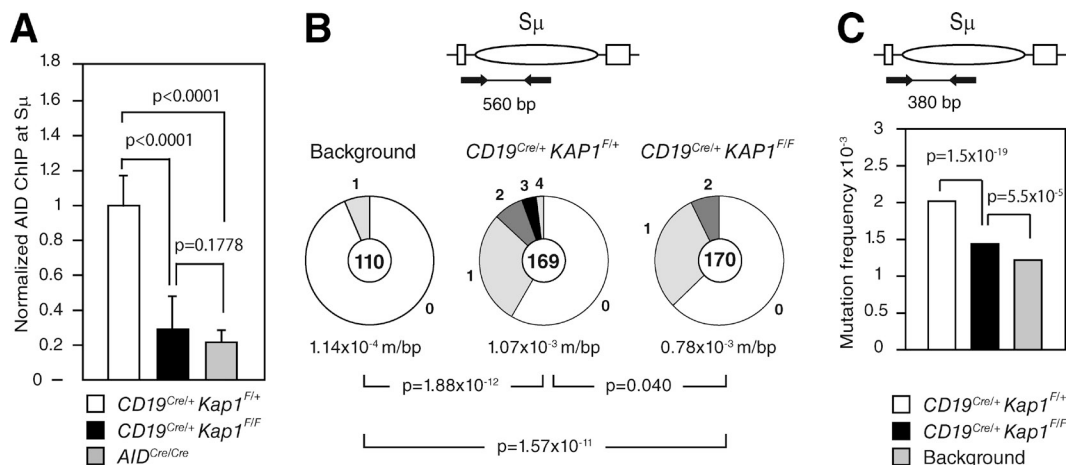


Figure 4. AID binding to S μ is impaired by *Kap1* deficiency and results in reduced levels of DNA damage. (A) ChIP analysis for AID occupancy at the S μ switch region in *CD19^{Cre/+}Kap1^{F/+}*, *CD19^{Cre/+}Kap1^{F/F}*, and *AID^{Cre/Cre}* B cells cultured in vitro with LPS + IL-4 for 60 h. Normalized AID-ChIP data from three experiments assayed with two different primer sets are shown. For each sample, AID-ChIP values were normalized to the input control and subtracted from the bead-only negative control. AID-ChIP signal in *CD19^{Cre/+}Kap1^{F/+}* B cells was assigned an arbitrary value of 1. P-values were determined by the one-tailed Student's *t* test. Error bars indicate SD. (B and C) Mutation analysis in the 5' end of S μ in *CD19^{Cre/+}Kap1^{F/F}* and *CD19^{Cre/+}Kap1^{F/+}* LPS + IL-4-stimulated B cells (72 h) as determined by Sanger sequencing (B) or HTS (C). Data are from five and four independent experiments, respectively. Background mutation frequency was determined on tail genomic DNA. Pie charts in B are as in Fig. 2. P-values were determined by the two-tailed Student's *t* test.

(Yan et al., 2007), we found no significant increase ($P = 0.9$) in the percentage of metaphases with abnormalities in *Kap1*-deficient B cells (Fig. 3 F and Table S1). Consistent with this, we found that *Kap1* disruption did not result in increased frequency of AID-dependent IgH/c-myc translocations (Fig. 3 G). We conclude that in the absence of KAP1, DSBs generated at switch regions are efficiently repaired and do not lead to global or IgH-specific genomic instability. Therefore, KAP1 functions upstream of the AID-induced DNA breaks that initiate CSR.

AID binding to S μ is impaired by *Kap1* deficiency

To determine whether AID recruitment to the S μ switch region is dependent on KAP1, we performed chromatin immunoprecipitation (ChIP) analysis using an anti-AID antibody (Pavri et al., 2010) on chromatin prepared from *CD19^{Cre/+}Kap1^{F/+}*, *CD19^{Cre/+}Kap1^{F/F}*, and *AID^{Cre/Cre}* activated B cells (Fig. 4 A). We found that AID occupancy at the S μ switch region was significantly reduced ($P < 0.0001$) in *CD19^{Cre/+}Kap1^{F/F}* when compared with *CD19^{Cre/+}Kap1^{F/+}* B cells (Fig. 4 A). AID-ChIP signal in *CD19^{Cre/+}Kap1^{F/F}* B cells relative to controls ranged between 20 and 50% (depending on the experiment; not depicted) and was specific, as only background levels of immunoprecipitated AID protein were observed at S μ in *AID^{Cre/Cre}* B cells. We conclude that AID binding to the donor switch region is impaired in *Kap1*-deficient B cells.

Impaired AID binding results in reduced levels of DNA damage at the S μ

To determine whether impaired AID recruitment in *Kap1*-deficient B cells translates into reduced levels of DNA damage, we measured the mutation frequency in the 5' end of S μ by cloning and sequencing. This assay has been used to assess

the level of AID-induced DNA damage in switch regions (Barreto et al., 2003; Guikema et al., 2007). We found that the proportion of mutated sequences and the mutation frequency were reduced ($P = 0.040$) in *CD19^{Cre/+}Kap1^{F/F}* ($F = 0.78 \times 10^{-3}$ mutations/bp; $n = 170$) when compared with controls ($F = 1.07 \times 10^{-3}$ mutations/bp; $n = 169$; Fig. 4 B). A significant reduction ($P = 1.5 \times 10^{-19}$) in mutation frequency in the 5' end of S μ was also found in *Kap1*-deficient B cells when sequences were analyzed in a larger dataset by high-throughput amplicon sequencing (HTS; Fig. 4 C). We conclude that impaired AID recruitment to S μ results in reduced levels of DNA damage and consequently in defective CSR.

AID forms a complex with KAP1 and HP1 that binds H3K9me3

Epigenetic modifications at the IgH locus have been suggested to target the CSR machinery to switch regions by relaxing chromatin or by providing binding motifs for factors involved in CSR (Odegard et al., 2005; Wang et al., 2006; Fraenkel et al., 2007; Chowdhury et al., 2008; Kuang et al., 2009; Wang et al., 2009). Among these, H3K9me3, which is usually associated with silent genes and heterochromatin, was found at transcribed donor and acceptor switch regions (Chowdhury et al., 2008; Kuang et al., 2009). We hypothesized that AID could be tethered to transcribed switch regions bearing this modification through the association between AID, KAP1, and HP1 and the ability of the latter to bind H3K9me3 (Bannister et al., 2001). To determine whether AID also associates with HP1, we performed reciprocal coimmunoprecipitation experiments using nuclear extracts prepared from cells expressing AID^{Flag-HA} and antibodies specific for the different HP1 isoforms. Although AID^{Flag-HA} coprecipitated with all HP1 isoforms, only HP1- γ reciprocally

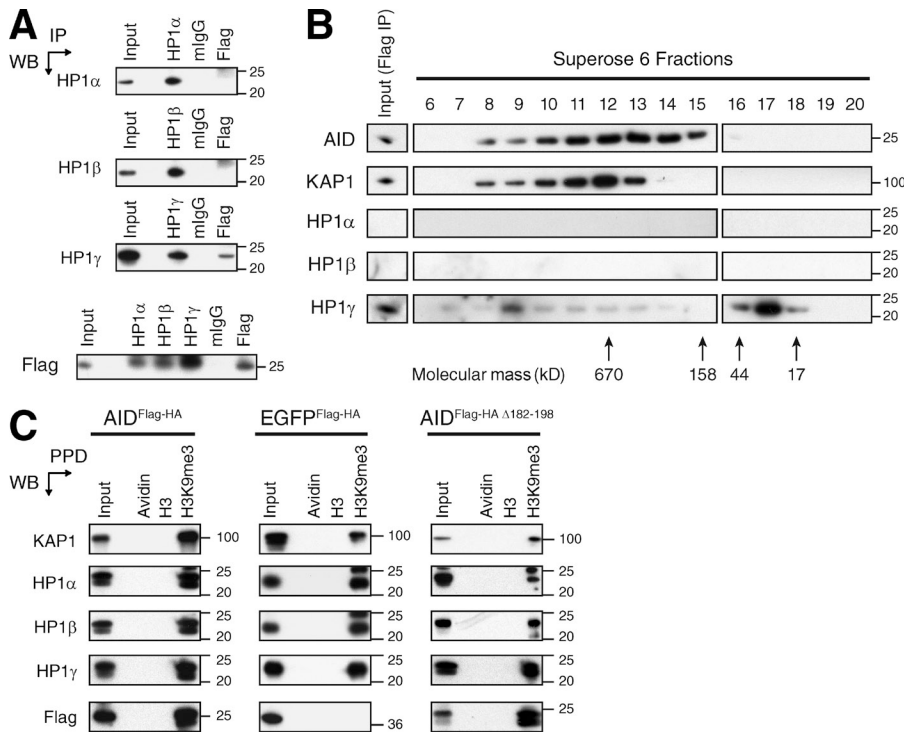


Figure 5. AID forms a complex with KAP1 and HP1 that binds H3K9me3.

(A) Nuclear extracts prepared from CH12 cells expressing AID^{Flag-HA} were immunoprecipitated and blotted with anti-HP1- α , anti-HP1- β , anti-HP1- γ , and anti-Flag antibodies. Data are representative of three independent experiments. (B) Flag immunoprecipitates eluted with the Flag peptide were fractionated with a Superose 6 gel filtration column. The indicated fractions were analyzed by Western blot using antibodies specific for KAP1, HP1- α , HP1- β , HP1- γ , and AID. Arrows indicate the elution position of calibration proteins of known molecular mass. Data are representative of two independent experiments. (C) Peptide pull-downs (PPD) using biotinylated unmodified (H3) or modified H3 peptides (H3K9me3) and avidin-agarose beads (Avidin) on nuclear extracts prepared from CH12 cells stably expressing AID^{Flag-HA}, EGFP^{Flag-HA}, or AID^{Flag-HA} Δ 182-198. Precipitated proteins were separated by SDS-PAGE and blotted with antibodies specific for KAP1, HP1- α , HP1- β , HP1- γ , and Flag. Data are representative of three independent experiments. (A-C) Molecular mass markers in kilodaltons are indicated. IP, immunoprecipitation; WB, Western blot.

coprecipitated AID^{Flag-HA} (Fig. 5 A). To determine whether KAP1, HP1, and AID exist in a complex, we fractionated Flag immunoprecipitates by gel filtration chromatography using a Superose 6 column and assayed the different fractions for AID, KAP1, and HP1 isoforms by Western blot (Fig. 5 B). We found that AID^{Flag-HA}, KAP1, and HP1- γ coelute in fractions corresponding to a complex of \sim 670 kD (Fig. 5 B). However, the majority of HP1- γ was eluted in low molecular weight fractions (Fig. 5 B), indicating that only a small fraction of the complex contains HP1- γ or that the *in vitro* association of HP1- γ within the complex is weaker. HP1- α and HP1- β were not detected (Fig. 5 B). To determine whether the AID-KAP1-HP1 complex is able to bind H3K9me3 residues, we performed peptide pull-down experiments using nuclear extracts from cells expressing AID^{Flag-HA}, EGFP^{Flag-HA}, or AID^{Flag-HA} Δ 182-198 and biotinylated histone H3 peptides either unmodified or trimethylated at lysine 9. We found that the H3K9me3 peptide precipitated KAP1, all the HP1 isoforms, and AID^{Flag-HA} (Fig. 5 C). Precipitation of AID^{Flag-HA} was specific, as this peptide did not precipitate EGFP^{Flag-HA} (Fig. 5 C). The H3K9me3 peptide also precipitated AID^{Flag-HA} Δ 182-198 (Fig. 5 C), indicating that the C-terminal domain of AID is not required for the recognition of H3K9me3 residues by the AID-KAP1-HP1 complex. We conclude that KAP1, HP1, and AID exist in a complex that binds H3K9me3 *in vitro*. Therefore, this modification could serve as a tethering motif for a complex formed by AID, KAP1, and HP1 during CSR *in vivo*.

H3K9me3 and KAP1 mark the donor switch region before and during CSR

To determine whether KAP1 and HP1 are recruited to switch regions during CSR *in vivo* and whether this correlates

with histone modifications associated with the IgH locus, we performed ChIP experiments on resting and activated control and *CD19^{Cre/+}Kap1^{F/F}* B cells (Fig. 6 and Figs. S6 and S7). Consistent with a previous study (Wang et al., 2009), we found that histone H3 associated with the pair of recombining switch regions is acetylated at lysine 9 and 14 (H3K9/K14Ac) in B cells undergoing CSR (Fig. 6 A) and that this correlated with switch region transcription (Fig. 3 A). Consistent with normal levels of germline transcripts in *Kap1*-deficient B cells (Fig. 3 A), we found that H3K9/K14Ac was not affected by *Kap1* deficiency (Fig. S6). Although H3K9me3 has been reported to be induced at the donor and acceptor switch regions during CSR (Chowdhury et al., 2008; Kuang et al., 2009), we found that H3K9me3 was already present at the IgH locus before recombination in resting B cells (Fig. 6 B). Upon stimulation, the level of H3K9me3 was substantially reduced throughout the locus but was specifically retained over the donor switch region (S μ ; Fig. 6 B and Table S2) when compared with sequences located immediately downstream of the variable region (J_H4), the I exons, the C_H exons, or regions surrounding the *Mest* promoter, which were included as positive and negative controls (Riclet et al., 2009). Surprisingly, H3K9me3 retention was not observed at S γ 3 when cells were stimulated to switch to IgG3 with LPS, nor at S γ 1 when cells were stimulated to switch to IgG1 with LPS + IL-4 (Fig. 6 B). The H3K9me3 pattern was similar in *Kap1*-deficient B cells (Fig. S6 and Table S3), indicating that this modification is not dependent on KAP1. Interestingly, we found that KAP1, together with

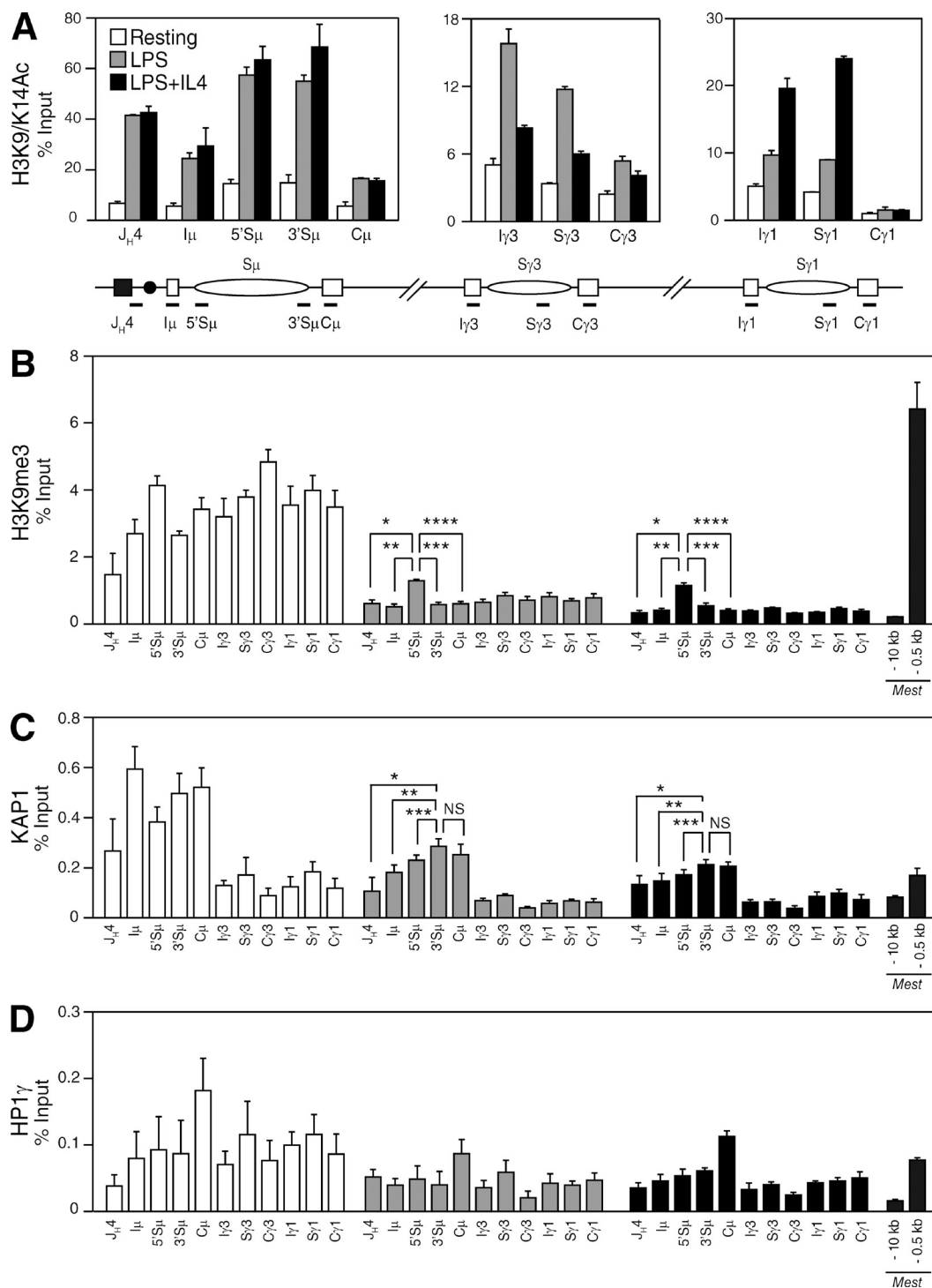


Figure 6. H3K9me3 and KAP1 mark the donor switch region (S_{μ}) before and during CSR. (A–D) ChIP analysis performed on chromatin prepared from resting or LPS-stimulated (48 h) and LPS + IL-4-stimulated (48 h) B cells obtained from *CD19^{Cre/+}Kap1^{fl/+}* mice using antibodies specific for H3K9/K14Ac (A), H3K9me3 (B), KAP1 (C), and HP1- γ (D). Real-time quantitative PCR was performed by using primer pairs specific for J_H4, I_μ, S_μ, C_μ, I_γ3, S_γ3, C_γ3, I_γ1, S_γ1, and C_γ1. Regions surrounding the *Mest* promoter (–10 kb and –0.5 kb) were used as negative and positive controls, respectively (Riclet et al., 2009). Fold change over control IgG is expressed as a percentage of input. Mean (+SD) of triplicate samples is shown. P-values were determined by the one-tailed Student's *t* test (H3K9me3, LPS: *, $P = 0.0003$; **, $P < 0.0001$; ***, $P < 0.0001$; ****, $P < 0.0001$; H3K9me3, LPS + IL-4: *, $P < 0.0001$; **, $P = 0.0001$; ***, $P = 0.0004$; ****, $P < 0.0001$; KAP1, LPS: *, $P = 0.0040$; **, $P = 0.0065$; ***, $P = 0.0289$; KAP1, LPS + IL-4: *, $P = 0.0134$; **, $P = 0.0170$; ***, $P = 0.0346$). See [Tables S2 and S5](#) for detailed statistical analysis and primer sequences, respectively. Data are representative of four independent experiments (see [Fig. S7](#) for an additional experiment).

H3K9me3, is recruited to the IgH locus in resting B cells. However, KAP1 binding was predominant over S μ when compared with downstream regions (Fig. 6 C). Upon stimulation, KAP1 was released from the locus but was

preferentially retained over S μ (Fig. 6 C). KAP1 recruitment was specific as no significant ChIP signal above background was found in *Kap1*-deficient B cells (not depicted). HP1- γ binding differed from H3K9me3 and KAP1 in that it was distributed throughout the locus in resting and activated B cells (Fig. 6 D, Fig. S7, and Table S4). Increased HP1- γ binding at C μ in activated B cells was observed (Fig. 6 D) but was not reproducible (Fig. S7). We conclude that H3K9me3, HP1- γ , and KAP1 are present at the IgH locus as part of a heterochromatic signature in resting B cells and that H3K9me3 and KAP1 specifically mark the donor switch region (S μ) before and during CSR.

The in vivo KAP1–HP1 association is required for efficient AID binding at S μ and CSR

To determine whether the in vivo association between KAP1 and HP1 is required to sustain CSR, we bred *CD19^{Cre/+} Kap1^{F/F}* mice with knockin mice (*Kap1^{V488A-L490A/F}*; Herzog et al., 2011) expressing a mutated KAP1 protein that is unable to associate with the HP1 proteins (Cammass et al., 2004) to generate *CD19^{Cre/+} Kap1^{V488A-L490A/F}* mice. Although B cells isolated from *CD19^{Cre/+} Kap1^{V488A-L490A/+}* mice (expressing a wild-type KAP1 and a mutant protein) underwent CSR at wild-type frequencies (similar to *CD19^{Cre/+} Kap1^{F/+}*), B cells isolated from *CD19^{Cre/+} Kap1^{V488A-L490A/F}* mice (expressing only the KAP1 mutant protein, which cannot associate with HP1; Fig. S2) displayed the same CSR defect observed in *CD19^{Cre/+} Kap1^{F/F}* B cells (Fig. 7, A and B). We conclude that the in vivo association between KAP1 and HP1 is required to sustain efficient CSR.

To determine whether AID binding to S μ is dependent on the interaction between KAP1 and HP1, we performed AID–ChIP experiments on chromatin prepared from *CD19^{Cre/+} Kap1^{F/+}*, *CD19^{Cre/+} Kap1^{V488A-L490A/F}*, and *AID^{Cre/Cre}* activated B cells (Fig. 7 C). We found that AID occupancy at the S μ switch region was significantly reduced ($P < 0.0001$) in *CD19^{Cre/+} Kap1^{V488A-L490A/F}* when compared with *CD19^{Cre/+} Kap1^{F/+}* B cells (Fig. 7 C). The reduction in AID retention at S μ was similar to that observed in *Kap1*-deficient B cells (Fig. 4 A). We conclude that the in vivo association between KAP1 and HP1 is required to mediate the efficient binding of AID to the donor switch region and to sustain CSR and that

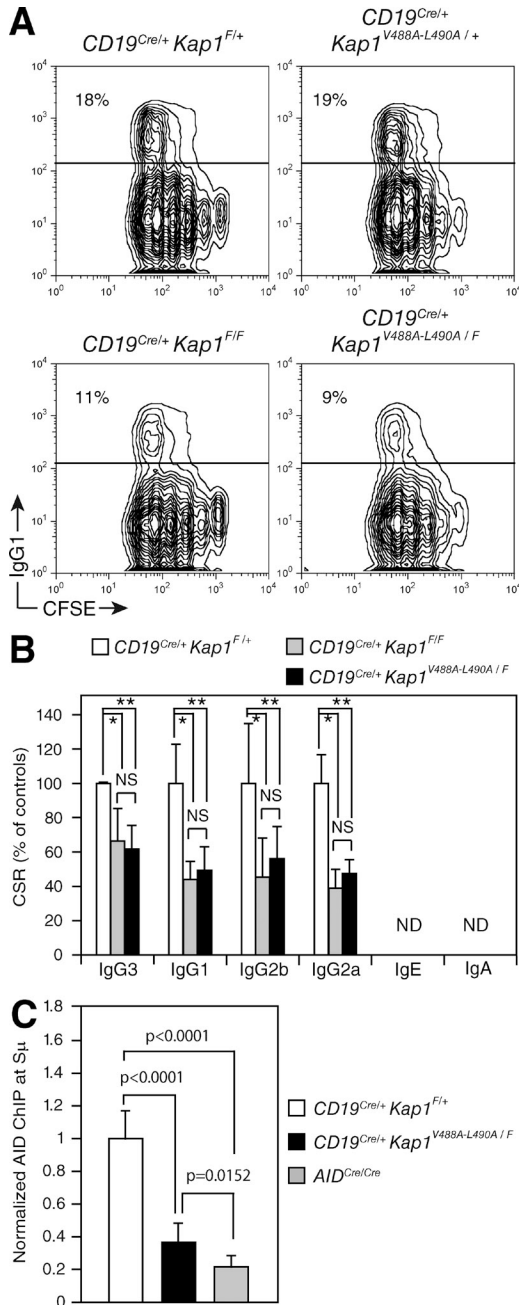


Figure 7. The in vivo association between KAP1 and HP1 is required for efficient CSR. (A) IgG1 cell surface expression in *CD19^{Cre/+} Kap1^{F/+}*, *CD19^{Cre/+} Kap1^{V488A-L490A/+}*, *CD19^{Cre/+} Kap1^{F/F}*, and *CD19^{Cre/+} Kap1^{V488A-L490A/F}* CFSE-labeled B cells stimulated with LPS + IL-4 for 3 d. The percentage of switched cells is indicated in each plot. (B) Percentage (+SD) of CSR in *CD19^{Cre/+} Kap1^{F/+}*, *CD19^{Cre/+} Kap1^{F/F}*, and *CD19^{Cre/+} Kap1^{V488A-L490A/F}* relative to *CD19^{Cre/+} Kap1^{F/+}*. CSR in *CD19^{Cre/+} Kap1^{F/+}* B cells was set to 100%. Data are from five independent experiments. P-values were

determined by the one-tailed Student's *t* test (*, *CD19^{Cre/+} Kap1^{F/+}* vs. *CD19^{Cre/+} Kap1^{F/F}*; IgG3, $P = 0.0154$; IgG1, $P = 0.0035$; IgG2b, $P = 0.0264$; IgG2a, $P = 0.001$; **, *CD19^{Cre/+} Kap1^{F/+}* vs. *CD19^{Cre/+} Kap1^{V488A-L490A/F}*; IgG3, $P = 0.0017$; IgG1, $P = 0.0036$; IgG2b, $P = 0.0276$; IgG2a, $P = 0.0004$). (C) ChIP analysis for AID occupancy at the S μ switch region in *CD19^{Cre/+} Kap1^{F/+}*, *CD19^{Cre/+} Kap1^{V488A-L490A/F}*, and *AID^{Cre/Cre}* B cells cultured in vitro with LPS + IL-4 for 60 h. Normalized AID–ChIP data from three experiments assayed with two different primer sets are shown. For each sample, AID–ChIP values were normalized to the input control and subtracted from the bead-only negative control. AID–ChIP signal in *CD19^{Cre/+} Kap1^{F/+}* B cells was assigned an arbitrary value of 1. P-values were determined by the one-tailed Student's *t* test. Error bars indicate SD.

a complex formed by AID, KAP1, and HP1 is tethered to the transcribed donor switch region bearing the H3K9me3 mark *in vivo*.

DISCUSSION

We have found that AID forms a complex with KAP1 and HP1 and that *Kap1* inactivation results in a CSR-specific defect that phenocopies mice deficient for components of the DDR (i.e., defective CSR and normal SHM; Ramiro et al., 2007). Because KAP1 has been reported to be an effector of the DDR downstream of ATM (Ziv et al., 2006) and because *Kap1* inactivation results in a CSR-specific defect similar to *Atm*^{-/-} mice (Lumsden et al., 2004; Reina-San-Martin et al., 2004), we hypothesized that the function of KAP1 in CSR is to respond to AID-induced DNA damage. However, we found that the role of KAP1 in CSR is independent of its phosphorylation status at serine 824 and that its inactivation does not result in defective DNA repair or increased genomic instability. Therefore, KAP1 does not appear to participate in the repair of AID-induced DNA damage during CSR.

Impaired CSR in the absence of KAP1 could be explained by a global reactivation of genes normally silenced by a KAP1-dependent mechanism (Nielsen et al., 1999) or by a substantial deregulation of chromatin structure (Ziv et al., 2006) at the IgH locus. However, this is unlikely as the transcriptional repressor activity of KAP1 requires its association with HP1 (Sripathy et al., 2006), and we show that the expression of a KAP1 mutant protein that cannot associate with HP1 (Cammass et al., 2004) was unable to restore CSR to wild-type levels. Furthermore, no significant alterations in switch region transcription were observed because of *Kap1* deficiency. We believe that the long-range interactions between switch region promoters and IgH enhancers, which are controlled by transcription (Wuerffel et al., 2007), are most likely not affected and that the global three-dimensional structure of the IgH locus is not altered by *Kap1* deficiency. The local access of AID to the switch region is most likely not affected in the absence of KAP1 because AID targeting to switch regions appears to be dependent on its association with Spt5 and RNA polymerase II (Pavri et al., 2010) and because *Kap1* inactivation reduces but does not abolish the recruitment of AID to S μ .

We pinpoint the role of KAP1 in CSR to a step that facilitates AID recruitment and the subsequent generation of DNA damage at S μ . We find a substantial reduction in AID recruitment that is in contrast with the 50% reduction in CSR that we observe in *Kap1*-deficient B cells. This might relate to differences in the immunoprecipitation efficiency between experiments. As the AID-ChIP signal in *Kap1*-deficient B cells (relative to controls) ranged between 20 and 50% depending on the experiment, we think that it might be underestimated.

We show that AID forms a complex with KAP1 and HP1 that selectively recognizes H3K9me3 *in vitro*. We also show that H3K9me3, KAP1, and HP1 are recruited to the IgH locus in resting B cells and that upon activation H3K9me3 and KAP1 are released but retained over S μ . With respect to the

pattern of H3K9me3, our results differ from a previous study (Kuang et al., 2009) as we find that this modification is not inducible. Although it is difficult to pinpoint the reason for this discrepancy, it most likely relates to technical issues such as differences in culture conditions, time points assayed, detection methods, etc. Nevertheless, the specific enrichment of KAP1 binding at S μ that we observe suggests that H3K9me3 can serve as a recognition motif for an AID-KAP1-HP1 complex *in vivo* at the donor switch region during CSR. Although we cannot provide an explanation as to why KAP1 recruitment expands up to C μ or why it is not recruited to the acceptor switch regions, this suggests that additional factors and/or histone modifications may be required to limit the range of AID recruitment and/or to modulate the efficiency of its retention. As we did not observe recruitment of KAP1 at S γ 3 or S γ 1, it is possible that KAP1 deficiency has no impact in the tethering efficiency of AID to acceptor switch regions, the efficiency of cytidine deamination, and the subsequent generation of DSBs. As DSB formation at acceptor switch regions has been suggested to be a rate-limiting step in CSR (Reina-San-Martin et al., 2003; Schrader et al., 2003), it might explain why the impairment in CSR in the absence of KAP1 is not more profound.

H3K9me3 is usually associated with heterochromatin formation and gene silencing. However, increasing evidence suggests additional roles for heterochromatin (Grewal and Jia, 2007), for example in recruiting factors that facilitate access of RNA polymerase II to heterochromatic regions (Zofall and Grewal, 2006) or by acting as a recruitment platform for factors regulating long-range chromosomal interactions (Jia et al., 2004). The presence of HP1- γ throughout the IgH locus before and during CSR is reminiscent of the mating-type switching in the fission yeast *Schizosaccharomyces pombe*, in which heterochromatin promotes the spreading of the recombination-promoting complex across the mating locus and imposes structural constraints that are important for the choice of the recombination donor site (Jia et al., 2004). Indeed, H3K9me3 has been shown to be present, along with HP1- γ , at actively transcribed genes in mammalian cells (Vakoc et al., 2005) and at recombining switch regions in mouse and human B cells (Chowdhury et al., 2008; Kuang et al., 2009). Strikingly, we demonstrate that the *in vivo* KAP1-HP1 association is required for efficient AID binding at S μ and to sustain CSR by expressing a KAP1 mutant that cannot associate with HP1 (in a *Kap1*-deficient background).

Our observations are consistent with a model in which AID forms a complex with KAP1 and HP1 that is tethered to transcribed switch regions bearing the H3K9me3 mark. There is precedent for a model based on the recognition of a modified histone, as recognition of hypermethylated histone H3 at lysine 4 by the PHD domain of RAG2 promotes efficient V(D)J recombination (Liu et al., 2007; Matthews et al., 2007). In our model, H3K9me3 and associated KAP1 and HP1- γ mark the S μ region before recombination. Upon stimulation, these marks are retained over the S μ and provide a docking motif for AID tethering *in vivo*. AID would then be retained

in close proximity to its substrate by KAP1, HP1, and H3K9me3, thus increasing the probability of cytidine deamination, the efficiency of DSB formation, and consequently of CSR. Our results thus provide a mechanism linking AID to epigenetic modifications associated with the IgH locus during CSR.

MATERIALS AND METHODS

Mice. *Kap1^{F/F}* (Cammass et al., 2000), *CD19^{Cre/+}* (provided by K. Rajewsky, Harvard Medical School, Boston, MA; Rickert et al., 1997), *Kap1^{V488A-LA90A/+}* (Herzog et al., 2011), and *AID^{Cre/Cre}* (Robbiani et al., 2008) mice were on a C57BL/6 background and were bred and maintained under specific pathogen-free conditions. Age-matched littermates (8–12 wk old) were used in all experiments. All animal work was performed under protocols approved by the Direction des Services Vétérinaires du Bas-Rhin, France (Authorization no. 67–343).

Splenic B cell purification, CSR assays, and two-color IgH-FISH.

Resting splenic B cells were isolated, cultured in vitro with LPS (CSR to IgG3 and IgG2b), LPS + IL-4 (CSR to IgG1 and IgE), or LPS + IFN- γ (CSR to IgG2a), and assayed for CSR as described previously (Robert et al., 2009). Switch transcripts, switch junctions, and IgH/c-myc translocations were analyzed as described previously (Robert et al., 2009). Metaphase spreads were prepared and subjected to two-color IgH-FISH as described previously (Franco et al., 2006).

Nuclear extracts and tandem affinity purification. Nuclear extracts and chromatin fractions were prepared from 12×10^9 cells using standard techniques. 20 mg of clarified extracts was taken into immunoprecipitation buffer (IP-300: 20 mM Tris, pH 7.9, 300 mM KCl, 20% glycerol, 0.25 mM EDTA, 0.125 mM EGTA, 0.5 mM PMSF, 1 mM DTT, 0.025% NP-40, 1 \times protease inhibitor cocktail [Roche], and 100 U/ml Benzonase [EMD]) and precleared with mouse IgG and protein G-agarose (GE Healthcare) for 1 h at 4°C. Flag M2 agarose beads (100 μ l, 50% slurry; Sigma-Aldrich) were added and incubated overnight at 4°C. Immune complexes were eluted three times with 100 μ l of 0.2 mg/ml Flag peptide (30 min at 4°C). Elutions were pooled and subjected to immunoprecipitation with anti-HA agarose beads (40 μ l, 50% slurry; 4 h at 4°C; Roche). Immune complexes were eluted twice with 40 μ l of 0.1 M glycine, pH 2.4 (10 min), or by overnight incubation with 40 μ l of 0.6 mg/ml HA peptide. Proteins were fractionated by SDS-PAGE in 4–12% gradient gels (Invitrogen) and stained with Coomassie blue (G250; Bio-Rad Laboratories) before processing for mass spectrometry.

Mass spectrometry analysis. Whole lanes from one-dimensional electrophoresis were systematically excised in \sim 5-mm bands. Proteins in bands were subjected to in-gel reduction, alkylation, and trypsin digestion. Peptides were dried, resuspended in 0.1% formic acid, and injected into the nano liquid chromatography (Ultimate 3000; Dionex) coupled with a linear ion trap mass spectrometer equipped with a nanoelectrospray source (LTQ XL ETD; Thermo Fisher Scientific). Tandem mass spectrometry spectra were recorded in the data-dependent mode on the five most intense ions observed in mass spectrometry scan with collision-induced dissociation and electron transfer dissociation fragmentation. Peptides selected for tandem mass spectrometry acquisition were then dynamically excluded for 30 s. Protein ID was performed using Proteome Discoverer 1.1 (Thermo Fisher Scientific) and SEQUEST (Thermo Fisher Scientific) searching against the mouse Swiss-Prot database (version 57.9) with the following fixed parameters: precursor mass tolerance of \pm 0.5 D, product ion mass tolerance of \pm 0.8 D, and two missed cleavages, and carbamidomethylation of cysteine as fixed modification and methionine oxidation as variable modification. SEQUEST results were filtered with Xcorr versus charge state 1.5–1, 1.6–2, 3–3, 3.2–4.

SHM analysis. Germinal center B cells (B220⁺Fas⁺GL-7⁺) were sorted from the lymph nodes of NP-CGG-immunized mice and analyzed for mutation in J_{H4} intron sequences (Jolly et al., 1997) with SHMTool (Maccarthy et al., 2009).

Real-time quantitative PCR. RNA and cDNA were prepared using standard techniques. Quantitative PCR was performed using SYBR green JumpStart Taq ReadyMix (Sigma-Aldrich) or QuantiTect SYBR green PCR kit (QIAGEN). Approximately 3 ng cDNA or 3–5 \times 10⁶ cell equivalents of de-cross-linked chromatin were run (in triplicate) and analyzed on a LightCycler 480 (Roche). Transcript or DNA quantities were calculated relative to standard curves and normalized to *CD79b* transcripts or input DNA. Gene of interest/*CD79b* or region of interest/input DNA ratios were averaged across experiments and normalized to the appropriate controls.

S μ mutation analysis. The 5' end of S μ was amplified by PCR using Pfu Turbo as described previously (Reina-San-Martin et al., 2004) or using Platinum Taq polymerase (Invitrogen) and bar-coded primers (Table S5) for HTS. HTS amplification conditions were as follows: 30 cycles of 94°C (15 s), 60°C (15 s), and 68°C (1 min). PCR products were purified with AmPure beads (Agencourt), quantified using Quan-iT Pico Green (Invitrogen), and mixed in equimolar ratios before sequencing with a 454 GS FLX sequencer (Roche; GATC). Sequences were aligned with Lasergene (DNASTAR) and analyzed with SHMTool (Maccarthy et al., 2009).

Peptide pull-downs. Avidin-conjugated agarose beads (Thermo Fisher Scientific) were coupled to biotinylated unmodified or modified H3 peptides (A₁RTKQTAR**K₆**[**me3**]STGGKAPRKQLATK₂₃-biotin; letters in bold indicate the modified amino acid, and the type of modification is in brackets) for 3 h at 25°C with rotation, washed in PBS 1 \times NP-40 0.1% and resuspended as a 50% slurry. 5 mg of nuclear extracts was precleared with avidin-agarose beads and incubated overnight with peptide-bound resin (40 μ l, 50% slurry). Resin was washed and resuspended in 20 μ l of Laemmli buffer. Proteins were fractionated by SDS-PAGE and analyzed by Western blot.

ChIP. The protocol was adapted from Millipore (<http://www.millipore.com/userguides/tech1/mcproto407>). In brief, 3 \times 10⁷ resting or stimulated B cells were cross-linked at 37°C for 10 min in 5 ml PBS/0.5% BSA with 1% formaldehyde. The reaction was quenched with 0.125 M glycine. After lysis, chromatin was sonicated to 0.5–1 kb using a Vibracell 75043 (Thermo Fisher Scientific). After 2 \times dilution in ChIP buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.1, and 167 mM NaCl), chromatin was precleared by rotating for 2 h at 4°C with 80 μ l of 50% protein A/G slurry (0.2 mg/ml sheared salmon sperm DNA, 0.5 mg/ml BSA, and 50% protein A/G; GE Healthcare). 0.3–0.5 \times 10⁶ cell equivalents were saved as input, and 3–5 \times 10⁶ cell equivalents were incubated overnight with specific or control antibodies (Table S6). Immune complexes were precipitated by the addition of protein A/G or Dynabeads protein A (for AID) and processed according to the Millipore protocol.

Gel filtration chromatography. 20 mg of nuclear extract was immunoprecipitated with Flag M2-agarose beads in IP-300 buffer, washed, and eluted with 0.2 mg/ml Flag peptide in 100 μ l. Eluted proteins were fractionated using a micro Superose 6 PC 3.2/30 column (ÅKTA). 100 μ l of proteins in each fraction was analyzed by Western blot.

Online supplemental material. Fig. S1 shows Western blot analysis of nuclear and cytoplasmic extracts prepared from CH12 cells expressing AID^{Flag-HA}, Western blot analysis of nuclear extracts from CH12 cells expressing AID^{Flag-HA}, EGFP^{Flag-HA}, and AID^{Flag-HA} Δ 182–198, SDS-PAGE (silver stain) analysis of proteins eluted after tandem affinity purification, and the KAP1 peptides identified by mass spectrometry. Fig. S2 shows that *Kap1* deficiency has no major impact on B cell development. Fig. S3 displays two additional experiments showing that *Kap1* deficiency results in reduced levels of post-recombination transcripts without decreasing germline transcription. Fig. S4 shows that *Kap1* deficiency has no effect on the frequency of germinal center B cells and the distribution or pattern of somatic mutations. Fig. S5 shows that AID overexpression in *Kap1*-deficient B cells does not rescue CSR. Fig. S6 displays two independent experiments showing that *Kap1* deficiency does not alter the pattern of H3K9me3 or H3K9/K14Ac marks at the IgH locus.

Fig. S7 displays an additional ChIP experiment (corresponding to Fig. 6) showing that H3K9me3 and KAP1 mark the donor switch region ($\Sigma\mu$) before and during CSR. Table S1 shows the two-color IgH-FISH analysis. Tables S2, S3, and S4, included as separate PDF files, show the statistical analyses corresponding to the ChIP experiments in Fig. 6, Fig. S6, and Fig. S7, respectively. Table S5 shows the primers used in this study. Table S6 shows the antibodies used in this study. Online supplemental material is available at <http://www.jem.org/cgi/content/full/jem.20110118/DC1>.

We thank members of the Losson and Reina-San-Martin laboratories for discussions and Evi Soutoglou and Olivier Pourquié for comments on the manuscript. We also thank Klaus Rajewsky for *CD19^{Cre}* mice, Michel Nussenzweig for the anti-AID antibody for ChIP analysis, Anna Gazumyan for advice on AID-ChIP, Raphaël Riclet for advice on H3K9me3, KAP1, and HP1 ChIP, Manuela Argenti and Frank Ruffenach for assistance in mass spectrometry analysis, Isabelle Billas for advice on gel filtration chromatography, Sonia Beroud for assistance with switch junction analysis, Claudine Ebel for assistance with cell sorting, Pascal Eberling for peptide synthesis, and Michael Gendron for animal care and immunizations.

B.P. Jeevan-Raj was supported by fellowships from the Association pour la Recherche sur le Cancer (ARC) and the Fondation pour la Recherche Médicale. F.W. Alt is an investigator of the Howard Hughes Medical Institute. This work was supported by grants to B. Reina-San-Martin from the Agence Nationale pour la Recherche (ANR-07-MIME-004-01), the ARC, and the Institut National de la Santé et de la Recherche Médicale (INSERM). B. Reina-San-Martin is an AVENIR-INSERM young investigator.

The authors have no conflicting financial interests.

Submitted: 17 January 2011

Accepted: 10 June 2011

REFERENCES

- Bannister, A.J., P. Zegerman, J.F. Partridge, E.A. Miska, J.O. Thomas, R.C. Allshire, and T. Kouzarides. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*. 410:120–124. doi:10.1038/35065138
- Barreto, V., B. Reina-San-Martin, A.R. Ramiro, K.M. McBride, and M.C. Nussenzweig. 2003. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol. Cell*. 12:501–508. doi:10.1016/S1097-2765(03)00309-5
- Brar, S.S., M. Watson, and M. Diaz. 2004. Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. *J. Biol. Chem.* 279:26395–26401. doi:10.1074/jbc.M403503200
- Cammas, F., M. Mark, P. Dollé, A. Dierich, P. Chambon, and R. Losson. 2000. Mice lacking the transcriptional corepressor TIF1beta are defective in early postimplantation development. *Development*. 127:2955–2963.
- Cammas, F., M. Herzog, T. Lerouge, P. Chambon, and R. Losson. 2004. Association of the transcriptional corepressor TIF1beta with heterochromatin protein 1 (HP1): an essential role for progression through differentiation. *Genes Dev.* 18:2147–2160. doi:10.1101/gad.302904
- Chaudhuri, J., U. Basu, A. Zarrin, C. Yan, S. Franco, T. Perlot, B. Vuong, J. Wang, R.T. Phan, A. Datta, et al. 2007. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv. Immunol.* 94:157–214. doi:10.1016/S0065-2776(06)94006-1
- Chowdhury, M., O. Forouhi, S. Dayal, N. McCloskey, H.J. Gould, G. Felsenfeld, and D.J. Fear. 2008. Analysis of intergenic transcription and histone modification across the human immunoglobulin heavy-chain locus. *Proc. Natl. Acad. Sci. USA*. 105:15872–15877. doi:10.1073/pnas.0808462105
- Di Noia, J.M., and M.S. Neuberger. 2007. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* 76:1–22. doi:10.1146/annurev.biochem.76.061705.090740
- Doi, T., L. Kato, S. Ito, R. Shinkura, M. Wei, H. Nagaoka, J. Wang, and T. Honjo. 2009. The C-terminal region of activation-induced cytidine deaminase is responsible for a recombination function other than DNA cleavage in class switch recombination. *Proc. Natl. Acad. Sci. USA*. 106:2758–2763. doi:10.1073/pnas.0813253106
- Fraenkel, S., R. Mostoslavsky, T.I. Novobrantseva, R. Pelanda, J. Chaudhuri, G. Esposito, S. Jung, F.W. Alt, K. Rajewsky, H. Cedar, and Y. Bergman. 2007. Allelic ‘choice’ governs somatic hypermutation in vivo at the immunoglobulin kappa-chain locus. *Nat. Immunol.* 8:715–722. doi:10.1038/ni1476
- Franco, S., M. Gostissa, S. Zha, D.B. Lombard, M.M. Murphy, A.A. Zarrin, C. Yan, S. Tepsuporn, J.C. Morales, M.M. Adams, et al. 2006. H2AX prevents DNA breaks from progressing to chromosome breaks and translocations. *Mol. Cell*. 21:201–214. doi:10.1016/j.molcel.2006.01.005
- Geisberger, R., C. Rada, and M.S. Neuberger. 2009. The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proc. Natl. Acad. Sci. USA*. 106:6736–6741. doi:10.1073/pnas.0810808106
- Grewal, S.I., and S. Jia. 2007. Heterochromatin revisited. *Nat. Rev. Genet.* 8:35–46. doi:10.1038/nrg2008
- Guikema, J.E., E.K. Linehan, D. Tschimoto, Y. Nakabeppu, P.R. Strauss, J. Stavnezer, and C.E. Schrader. 2007. APE1- and APE2-dependent DNA breaks in immunoglobulin class switch recombination. *J. Exp. Med.* 204:3017–3026. doi:10.1084/jem.20071289
- Herzog, M., O. Wendling, F. Guillou, P. Chambon, M. Mark, R. Losson, and F. Cammas. 2011. TIF1β association with HP1 is essential for post-gastrulation development, but not for Sertoli cell functions during spermatogenesis. *Dev. Biol.* 350:548–558. doi:10.1016/j.ydbio.2010.12.014
- Ito, S., H. Nagaoka, R. Shinkura, N. Begum, M. Muramatsu, M. Nakata, and T. Honjo. 2004. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl. Acad. Sci. USA*. 101:1975–1980. doi:10.1073/pnas.0307335101
- Jia, S., T. Yamada, and S.I. Grewal. 2004. Heterochromatin regulates cell type-specific long-range chromatin interactions essential for directed recombination. *Cell*. 119:469–480. doi:10.1016/j.cell.2004.10.020
- Jolly, C.J., N. Klix, and M.S. Neuberger. 1997. Rapid methods for the analysis of immunoglobulin gene hypermutation: application to transgenic and gene targeted mice. *Nucleic Acids Res.* 25:1913–1919. doi:10.1093/nar/25.10.1913
- Kracker, S., P. Gardes, F. Mazerolles, and A. Durandy. 2010. Immunoglobulin class switch recombination deficiencies. *Clin. Immunol.* 135:193–203. doi:10.1016/j.clim.2010.01.012
- Kuang, F.L., Z. Luo, and M.D. Scharff. 2009. H3 trimethyl K9 and H3 acetyl K9 chromatin modifications are associated with class switch recombination. *Proc. Natl. Acad. Sci. USA*. 106:5288–5293. doi:10.1073/pnas.0901368106
- Liu, Y., R. Subrahmanyam, T. Chakraborty, R. Sen, and S. Desiderio. 2007. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity*. 27:561–571. doi:10.1016/j.immuni.2007.09.005
- Lumsden, J.M., T. McCarty, L.K. Petiot, R. Shen, C. Barlow, T.A. Wynn, H.C. Morse III, P.J. Gearhart, A. Wynshaw-Boris, E.E. Max, and R.J. Hodes. 2004. Immunoglobulin class switch recombination is impaired in *Atm*-deficient mice. *J. Exp. Med.* 200:1111–1121. doi:10.1084/jem.20041074
- Maccarthy, T., S. Roa, M.D. Scharff, and A. Bergman. 2009. SHMTool: a webserver for comparative analysis of somatic hypermutation datasets. *DNA Repair (Amst.)*. 8:137–141. doi:10.1016/j.dnarep.2008.09.006
- Mathews, A.G., A.J. Kuo, S. Ramón-Maiques, S. Han, K.S. Champagne, D. Ivanov, M. Gallardo, D. Carney, P. Cheung, D.N. Ciccone, et al. 2007. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature*. 450:1106–1110. doi:10.1038/nature06431
- McBride, K.M., V. Barreto, A.R. Ramiro, P. Stavropoulos, and M.C. Nussenzweig. 2004. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J. Exp. Med.* 199:1235–1244. doi:10.1084/jem.20040373
- Muramatsu, M., K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai, and T. Honjo. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*. 102:553–563. doi:10.1016/S0092-8674(00)00078-7
- Nielsen, A.L., J.A. Ortiz, J. You, M. Oulad-Abdelghani, R. Khechumian, A. Gansmuller, P. Chambon, and R. Losson. 1999. Interaction with

- members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *EMBO J.* 18:6385–6395. doi:10.1093/emboj/18.22.6385
- Nonaka, T., T. Doi, T. Toyoshima, M. Muramatsu, T. Honjo, and K. Kinoshita. 2009. Carboxy-terminal domain of AID required for its mRNA complex formation in vivo. *Proc. Natl. Acad. Sci. USA.* 106:2747–2751. doi:10.1073/pnas.0812957106
- Odegard, V.H., S.T. Kim, S.M. Anderson, M.J. Shlomchik, and D.G. Schatz. 2005. Histone modifications associated with somatic hypermutation. *Immunity.* 23:101–110. doi:10.1016/j.immuni.2005.05.007
- Pavri, R., A. Gazumyan, M. Jankovic, M. Di Virgilio, I. Klein, C. Ansarah-Sobrinho, W. Resch, A. Yamane, B. Reina San-Martin, V. Barreto, et al. 2010. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Sp5. *Cell.* 143:122–133. doi:10.1016/j.cell.2010.09.017
- Petersen-Mahrt, S.K., R.S. Harris, and M.S. Neuberger. 2002. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature.* 418:99–103. doi:10.1038/nature00862
- Ramiro, A., B. Reina San-Martin, K. McBride, M. Jankovic, V. Barreto, A. Nussenzweig, and M.C. Nussenzweig. 2007. The role of activation-induced deaminase in antibody diversification and chromosome translocations. *Adv. Immunol.* 94:75–107. doi:10.1016/S0065-2776(06)94003-6
- Reina-San-Martin, B., S. Difilippantonio, L. Hanitsch, R.F. Masilamani, A. Nussenzweig, and M.C. Nussenzweig. 2003. H2AX is required for recombination between immunoglobulin switch regions but not for intra-switch region recombination or somatic hypermutation. *J. Exp. Med.* 197:1767–1778. doi:10.1084/jem.20030569
- Reina-San-Martin, B., H.T. Chen, A. Nussenzweig, and M.C. Nussenzweig. 2004. ATM is required for efficient recombination between immunoglobulin switch regions. *J. Exp. Med.* 200:1103–1110. doi:10.1084/jem.20041162
- Rickert, R.C., J. Roes, and K. Rajewsky. 1997. B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucleic Acids Res.* 25:1317–1318. doi:10.1093/nar/25.6.1317
- Riclet, R., M. Chendeb, J.L. Vonesch, D. Koczan, H.J. Thiesen, R. Losson, and F. Cammas. 2009. Disruption of the interaction between transcriptional intermediary factor 1beta and heterochromatin protein 1 leads to a switch from DNA hyper- to hypomethylation and H3K9 to H3K27 trimethylation on the MEST promoter correlating with gene reactivation. *Mol. Biol. Cell.* 20:296–305. doi:10.1091/mbc.E08-05-0510
- Robbiani, D.F., A. Bothmer, E. Callen, B. Reina-San-Martin, Y. Dorsett, S. Difilippantonio, D.J. Bolland, H.T. Chen, A.E. Corcoran, A. Nussenzweig, and M.C. Nussenzweig. 2008. AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. *Cell.* 135:1028–1038. doi:10.1016/j.cell.2008.09.062
- Robert, I., F. Dantzer, and B. Reina-San-Martin. 2009. Parp1 facilitates alternative NHEJ, whereas Parp2 suppresses IgH/c-myc translocations during immunoglobulin class switch recombination. *J. Exp. Med.* 206:1047–1056. doi:10.1084/jem.20082468
- Schrader, C.E., S.P. Bradley, J. Vardo, S.N. Mochegova, E. Flanagan, and J. Stavnez. 2003. Mutations occur in the Ig Smu region but rarely in Sgamma regions prior to class switch recombination. *EMBO J.* 22:5893–5903. doi:10.1093/emboj/cdg550
- Sellers, M., B. Reina-San-Martin, P. Kastner, and S. Chan. 2009. Ikaros controls isotype selection during immunoglobulin class switch recombination. *J. Exp. Med.* 206:1073–1087. doi:10.1084/jem.20082311
- Sernández, I.V., V.G. de Yébenes, Y. Dorsett, and A.R. Ramiro. 2008. Haploinsufficiency of activation-induced deaminase for antibody diversification and chromosome translocations both in vitro and in vivo. *PLoS ONE.* 3:e3927. doi:10.1371/journal.pone.0003927
- Shinkura, R., S. Ito, N.A. Begum, H. Nagaoka, M. Muramatsu, K. Kinoshita, Y. Sakakibara, H. Hijikata, and T. Honjo. 2004. Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat. Immunol.* 5:707–712. doi:10.1038/ni1086
- Sripathy, S.P., J. Stevens, and D.C. Schultz. 2006. The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Mol. Cell. Biol.* 26:8623–8638. doi:10.1128/MCB.00487-06
- Ta, V.T., H. Nagaoka, N. Catalan, A. Durandy, A. Fischer, K. Imai, S. Nonoyama, J. Tashiro, M. Ikegawa, S. Ito, et al. 2003. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat. Immunol.* 4:843–848. doi:10.1038/ni964
- Vakoc, C.R., S.A. Mandat, B.A. Olenchok, and G.A. Blobel. 2005. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol. Cell.* 19:381–391. doi:10.1016/j.molcel.2005.06.011
- Vuong, B.Q., M. Lee, S. Kabir, C. Irimia, S. Macchiarulo, G.S. McKnight, and J. Chaudhuri. 2009. Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. *Nat. Immunol.* 10:420–426. doi:10.1038/ni.1708
- Wang, L., N. Whang, R. Wuerffel, and A.L. Kenter. 2006. AID-dependent histone acetylation is detected in immunoglobulin S regions. *J. Exp. Med.* 203:215–226. doi:10.1084/jem.20051774
- Wang, L., R. Wuerffel, S. Feldman, A.A. Khamlichi, and A.L. Kenter. 2009. S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. *J. Exp. Med.* 206:1817–1830. doi:10.1084/jem.20081678
- Ward, I.M., B. Reina-San-Martin, A. Oлару, K. Minn, K. Tamada, J.S. Lau, M. Cascalho, L. Chen, A. Nussenzweig, F. Livak, et al. 2004. 53BP1 is required for class switch recombination. *J. Cell Biol.* 165:459–464. doi:10.1083/jcb.200403021
- Wei, M., R. Shinkura, Y. Doi, M. Maruya, S. Fagarasan, and T. Honjo. 2011. Mice carrying a knock-in mutation of Aicda resulting in a defect in somatic hypermutation have impaired gut homeostasis and compromised mucosal defense. *Nat. Immunol.* 12:264–270. doi:10.1038/ni.1991
- Wu, X., P. Geraldes, J.L. Platt, and M. Cascalho. 2005. The double-edged sword of activation-induced cytidine deaminase. *J. Immunol.* 174:934–941.
- Wuerffel, R., L. Wang, F. Grigera, J. Manis, E. Selsing, T. Perlot, F.W. Alt, M. Cogne, E. Pinaud, and A.L. Kenter. 2007. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity.* 27:711–722. doi:10.1016/j.immuni.2007.09.007
- Xu, Z., Z. Fulop, G. Wu, E.J. Pone, J. Zhang, T. Mai, L.M. Thomas, A. Al-Qahtani, C.A. White, S.R. Park, et al. 2010. 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. *Nat. Struct. Mol. Biol.* 17:1124–1135. doi:10.1038/nsmb.1884
- Yamane, A., W. Resch, N. Kuo, S. Kuchen, Z. Li, H.W. Sun, D.F. Robbiani, K. McBride, M.C. Nussenzweig, and R. Casellas. 2011. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat. Immunol.* 12:62–69. doi:10.1038/ni.1964
- Yan, C.T., C. Boboila, E.K. Souza, S. Franco, T.R. Hickernell, M. Murphy, S. Gumaste, M. Geyer, A.A. Zarrin, J.P. Manis, et al. 2007. IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature.* 449:478–482. doi:10.1038/nature06020
- Ziv, Y., D. Bielopolski, Y. Galanty, C. Lukas, Y. Taya, D.C. Schultz, J. Lukas, S. Bekker-Jensen, J. Bartek, and Y. Shiloh. 2006. Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nat. Cell Biol.* 8:870–876. doi:10.1038/ncb1446
- Zofall, M., and S.I. Grewal. 2006. Swi6/HP1 recruits a JmjC domain protein to facilitate transcription of heterochromatic repeats. *Mol. Cell.* 22:681–692. doi:10.1016/j.molcel.2006.05.010

A role for the RNA pol II-associated PAF complex in AID-induced immune diversification

Katharina L. Willmann,^{1,2} Sara Milosevic,⁴ Siim Pauklin,¹ Kerstin-Maike Schmitz,² Gopinath Rangam,^{1,2} Maria T. Simon,¹ Sarah Maslen,³ Mark Skehel,³ Isabelle Robert,⁴ Vincent Heyer,⁴ Ebe Schiavo,⁴ Bernardo Reina-San-Martin,⁴ Svend K. Petersen-Mahrt^{1,2}

¹DNA Editing Laboratory, London Research Institute, South Mimms EN6 3LD, England, UK

²DNA Editing in Immunity and Epigenetics, IFOM-Fondazione Instituto FIRC di Oncologia Molecolare, Via Adamello 16, 20139 Milano, Italy

³Protein Analysis and Proteomics Laboratory, London Research Institute, South Mimms EN6 3LD, England, UK

⁴Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de la Santé et de la Recherche Médicale (INSERM) U964, Centre National de la Recherche Scientifique (CNRS) UMR7104, Université de Strasbourg, 67404 Illkirch, France

Antibody diversification requires the DNA deaminase AID to induce DNA instability at immunoglobulin (Ig) loci upon B cell stimulation. For efficient cytosine deamination, AID requires single-stranded DNA and needs to gain access to Ig loci, with RNA pol II transcription possibly providing both aspects. To understand these mechanisms, we isolated and characterized endogenous AID-containing protein complexes from the chromatin of diversifying B cells. The majority of proteins associated with AID belonged to RNA polymerase II elongation and chromatin modification complexes. Besides the two core polymerase subunits, members of the PAF complex, SUPT5H, SUPT6H, and FACT complex associated with AID. We show that AID associates with RNA polymerase-associated factor 1 (PAF1) through its N-terminal domain, that depletion of PAF complex members inhibits AID-induced immune diversification, and that the PAF complex can serve as a binding platform for AID on chromatin. A model is emerging of how RNA polymerase II elongation and pausing induce and resolve AID lesions.

CORRESPONDENCE

Svend K. Petersen-Mahrt:
svend.petersen-mahrt@ifom.eu

Abbreviations used: AID, activation-induced deaminase; ChIP, chromatin IP; CSR, class switch recombination; CTD, C-terminal domain (of RNA pol II); iGC, Ig gene conversion; IP, immunoprecipitation; PAF, RNA polymerase-associated factor; qRT-PCR, quantitative RT-PCR; SEC, size exclusion chromatography; SHM, somatic hypermutation; TSS, transcription start site.

In B cells, antibody diversity is created via two DNA instability mechanisms (Rajewsky, 1996). In the first, RAG1/2 mediate antigen-independent V(D)J recombination, and in the second, activation-induced deaminase (AID) drives antigen-dependent Ig diversification. The latter includes somatic hypermutation (SHM), Ig gene conversion (iGC), and class switch recombination (CSR). SHM and iGC induce variable (V) region diversification via templated and nontemplated DNA mutations (Di Noia and Neuberger, 2007), whereas CSR recombines DNA constant (C) switch regions, resulting in IgM to IgG, IgA, or IgE isotype switching (Stavnezer et al., 2008). Mechanistically, SHM,

iGC, and CSR are initiated by the DNA deaminase AID, which deaminates cytosine (dC) residues to uracil (dU) on single-stranded DNA (ssDNA; Petersen-Mahrt, 2002, 2005; Bransteitter et al., 2003; Chaudhuri et al., 2003). At the genetic level, deamination causes a change in base recognition, as uracil is read as thymine during replication. At the biochemical level, reformation of double-stranded DNA (dsDNA) causes an alteration of DNA structure, resulting in a dU:dG lesion, which in turn activates DNA repair pathways resulting in mutated or otherwise altered chromosomes.

Because of the high oncogenic potential of AID, understanding how DNA deaminases are

K.L. Willmann, S. Milosevic, and S. Pauklin contributed equally to this paper.

S. Pauklin's present address is Laboratory For Regenerative Medicine, University of Cambridge, Cambridge CB2 0SZ, England, UK.

© 2012 Willmann et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

regulated at the target site is one of the most important aspects in the field of DNA editing and Ig diversification; however, little is known about the protein complexes and mechanisms involved. Mechanistically, AID requires ssDNA as a substrate, and although several chromatin alteration events could lead to ssDNA formation, transcription at the Ig locus is required for SHM and CSR. The rate of transcription correlates with the rate of SHM (Peters and Storb, 1996), and germline transcription through the switch and the constant region precedes CSR (Stavnezer-Nordgren and Sirlin, 1986). Interaction of AID with CTNBL1 (Conticello et al., 2008; Ganesh et al., 2011) demonstrated an association with RNA processing. More recently, though, direct links between AID and mRNA transcription were demonstrated. It was shown that CSR required the basal transcription factor SUPT5H (Pavri et al., 2010) and its associated factor SUPT4H (Stanlie et al., 2012), the transcription-associated chromatin modifier FACT complex (Stanlie et al., 2010), and histone chaperon SUPT6H (Okazaki et al., 2011), whereas AID activity during CSR was enhanced by components of the RNA-processing exosome (Basu et al., 2011).

To delineate the biochemical link of RNA pol II transcription to AID-induced Ig diversification, and to further characterize the AID interactome, we developed a novel biochemical approach: we C-terminally tagged the endogenous AID protein in Ig diversifying cells with a FLAG or a FLAG/Myc epitope (Pauklin et al., 2009), and we adapted a recently developed method for isolation of chromatin-bound protein complexes (Aygün et al., 2008). This method allowed, for the first time, the identification and characterization of proteins that are associated with AID on chromatin in their physiological environment. The majority of the identified proteins (FACT complex, SUPT5H, SUPT6H, RNA polymerase-associated factor (PAF) complex, RPB1A, RPB1B, and DNA topo I) are involved in RNA processing, chromatin remodeling, exosome processing, and RNA pol II transcription elongation/pausing. We identified a direct interaction of AID (the N-terminal domain) with PAF1, and by using knockdown experiments, we could demonstrate physiological importance of the PAF complex for Ig class switching and recruitment of AID to the Ig locus. A model of how this complex could influence AID efficacy at the Ig locus will be discussed.

RESULTS

To determine the composition of the protein complexes that interact with AID on chromatin in B cells undergoing Ig receptor diversification, we developed cell lines in which endogenous AID was tagged with epitope-peptides at the C terminus (Pauklin et al., 2009). In the chicken B cell lymphoma DT40, which continuously undergoes AID-dependent diversification of the Ig locus, AID was tagged with either 3xFLAG peptides (3F) or the combination of 3xFLAG peptide, 2xTEV cleavage sites, and 3xMyc peptides (3FM). This yielded expression of tagged AID to levels that were comparable to endogenous amounts. Although it is known that the C terminus of AID plays an important

role in subcellular localization, we could not detect a significant change in AID relocalization or immune diversification activity caused by the monoallelic C-terminal tags (unpublished data).

Chromatin AID is part of a multimeric complex

Because AID is predominantly localized in the cytoplasm (Rada et al., 2002; Brar et al., 2004; Ito et al., 2004; McBride et al., 2004), and only limited amounts can be identified within the nucleus, we grew $1\text{--}2 \times 10^{10}$ AID-3FM cells for biochemical analysis. Cell lysates were subfractionated into cytoplasm, nucleoplasm, and chromatin fractions. We focused on chromatin-bound AID, which we estimated to be <2% of total AID-3FM (unpublished data). The isolated chromatin fraction was further separated using a Superdex 200 column for size exclusion chromatography (SEC), thereby determining the size of the AID-associated protein complex bound to chromatin (Fig. 1 a). AID was identified as part of a 200-kD protein complex (120–300 kD based on standard proteins), whereas only a minor fraction of AID eluted at its theoretical monomeric size of 27 kD (Fig. 1 a). This demonstrated that AID isolated from chromatin under physiological conditions is part of a large heteromeric complex.

The PAF complex associates with AID on chromatin in Ig diversifying cells

To identify proteins associated with chromatin-bound AID, we performed FLAG peptide immunoprecipitations (IPs), followed by one-dimension SDS-PAGE and mass spectrometry identification (Fig. S1). We obtained 1,319 peptide identities (Ids), corresponding to 391 proteins from AID-3FM cells. Mass spectrometric analysis of IPs from cytoplasmic, nucleoplasmic, and chromatin fractions of a control cell line (expressing AID without a tag) served as a control peptide Id database. Using this database, we eliminated 366 of the 391 proteins (>15-fold enrichment; all AID-interacting proteins are listed in Fig. S1). When we submitted the protein Ids into the Ingenuity Systems Pathway Analysis gene network software, we obtained a potential interacting network containing >80% of all isolated peptides (Fig. S2). The majority of the AID-associated proteins from the chromatin fraction were part of mRNA processing. Aside from the core RNA pol II subunits, we identified the core PAF complex (RNA polymerase-associated factor; PAF1, CTR9, LEO1), FACT complex (SSRP1, SUPT16H), SUPT5H, SUPT6H, and DNA topo I (Fig. 1 b). These factors play a direct role in RNA pol II pausing/restarting and elongation, as well as in chromatin modification and exosome processing. Furthermore, an additional 20 peptides comprised proteins involved in RNA metabolism (splicing-associated factors and RNA helicase). The high percentage (54%) of peptides that are part of the same biological process (early mRNA biogenesis), and which co-isolate with AID, indicated that our isolation and analysis procedure had identified key AID-interacting proteins at the chromatin level

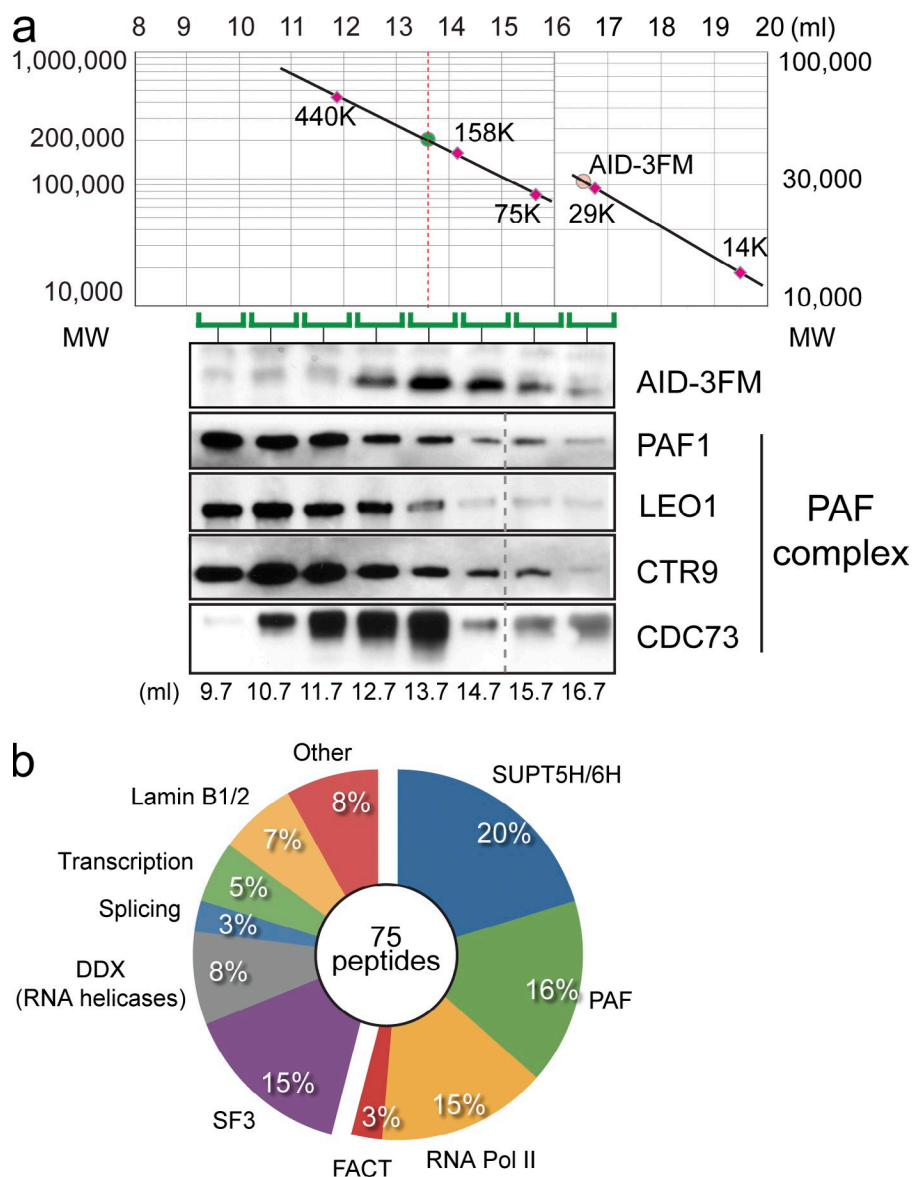


Figure 1. Chromatin-bound AID is in a multimeric complex. (a) Elution profile of size exclusion chromatography. The chromatin fraction from DT40 AID-3FM cells was loaded onto a Superdex 200 column. (top) The elution profile of standard proteins is plotted in the graph. Red circle, theoretical elution volume of AID-3FM (29 kD). (bottom) Eluted fractions (elution volume indicated, milliliters) were analyzed by Western blot probing for the presence of AID-3FM, PAF1, LEO1, CTR9, and CDC73. (b) Analysis of filtered peptide hits. The most abundant peptides identified were grouped using gene network software analysis, and groups are displayed in a pie chart as percentage of total peptides (75) that could be assigned to each complex. A full table of chromatin peptides is listed in Fig. S1.

or parafibromin), as well as confirm SUPT5H and SUPT6H (Fig. 2 a). The AID-association with SUPT5H (Fig. 2 a), although technically difficult, was further confirmed by multiple large scale FLAG immunoprecipitation and mass spectrometry, in which SUPT5H association was identified in three out of five experiments (and SUPT6H and PAF1 were identified in each IP). In conclusion, our work has for the first time identified and verified AID-associated complexes on chromatin in diversifying B cells.

The PAF complex associates with AID in CSR-competent murine B cells

To determine whether the identified associations between AID and RNA pol II-associated factors observed in DT40 cells is also present in murine

CSR-proficient cells, we performed a coIP experiment from nuclear extracts of CH12 B cells expressing tagged AID (AID^{Flag-HA}; Jeevan-Raj et al., 2011). Consistent with the DT40 analysis (Fig. 2 a), PAF1, LEO1, CTR9, SUPT5H, and SUPT6H could be identified to associate with AID (Fig. 2 b). Moreover in a reciprocal experiment, in which PAF1, LEO1, CTR9, SUPT5H, SUPT6H, and RNA polymerase II were precipitated, we identified AID in all IPs performed (unpublished data). This indicated that the identified AID associations were present in both DT40 and CH12 cells, thereby establishing a potential biochemical link between V region diversification (DT40 cells) and CSR (CH12 cells).

AID associates with the PAF complex via PAF1

To further characterize the PAF complex association with AID, we used immunoblot analysis of the chromatin SEC

from immune-diversifying cells. Consistent with this, several of the proteins that we identified (RNA pol II, SUPT5H, SUPT6H, FACT complex, and DNA topo I) have been previously described to play a role in SHM and CSR. It is important to note that the chicken genome is not fully characterized and annotated, and thus the number of proteins we have identified may be underestimated.

Our mass spectrometry analysis of the AID chromatin interactome showed PAF1, CTR9, and LEO1 as AID-associated proteins on chromatin. They form part of the PAF complex, a RNA pol II-associated complex that promotes elongation (Kim et al., 2010) by recruiting enzymes for histone H2 monoubiquitination and other co-transcriptional chromatin marks (Jaehning, 2010). We could verify the associations of AID by analyzing the chromatin FLAG-IP for PAF1 (two different antibodies), LEO1, CDC73 (also known as HRPT2

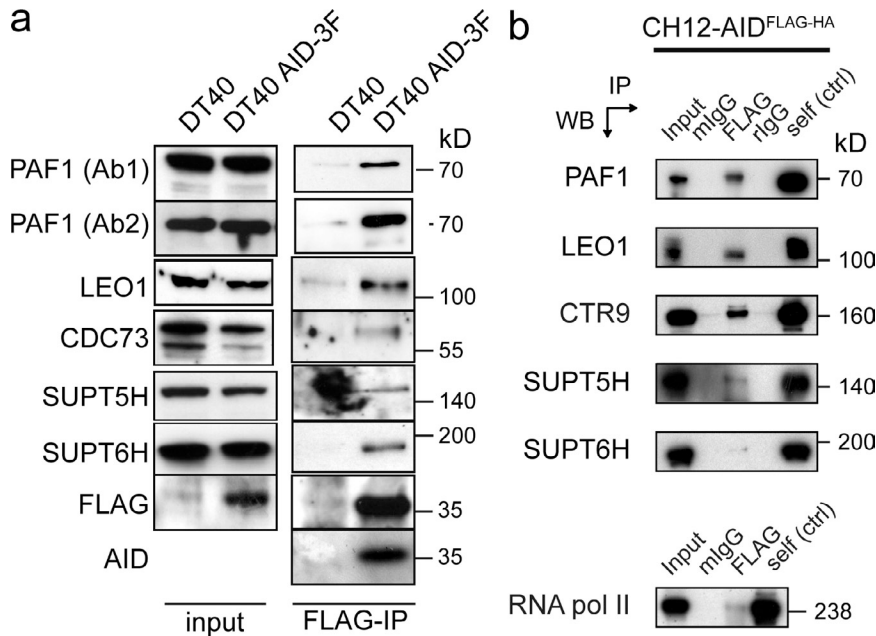


Figure 2. PAF complex and RNA pol II-associated proteins coIP with AID in Ig diversifying cells. (a) Verification of DT40 chromatin fractions were immunoprecipitated via FLAG and eluted complexes were probed by Western-blotting: anti-PAF1 (two antibodies), anti-LEO1, anti-CDC73, anti-SUPT5H, anti-SUPT6H, anti-FLAG, and anti-AID antibodies. Input lysates from both cell lines are shown on the left (input); parental DT40 cell line (DT40) served as negative control for FLAG-IPs from chromatin fractions. The anti-AID antiserum was unable to detect chicken AID at physiological levels (not shown). (b) Nuclear extracts from CH12 cells stably expressing AID^{FLAG-HA} were immunopurified via anti-FLAG and the following controls: murine IgGs (mIgG) and rabbit IgG (rIgG). The isolated complex was probed with the indicated antibodies. Input represents 1% of material used for IP. Self (ctrl) refers to an IP using the indicated antibody on the left. ND, not determined. Molecular weight in kilodaltons are indicated. Data are representative of three independent experiments.

fractions (Fig. 1 a), and demonstrated that PAF1, LEO1, and CTR9 co-migrate in a large (>400-kD) complex (Fig. 1 a, lanes 1–3), with the peak trailing fractions overlapping with the AID peak (Fig. 1 a, lane 5). Although AID did not fully co-migrate in the same peaks, the data indicated that the classical PAF complex was present in DT40 and partially associated with AID on chromatin. It was therefore likely that AID interacted with one of the components of the PAF complex rather than with each individual member.

We coexpressed AID with individual PAF members in *E. coli* and monitored binding by coIP and Western blot analysis (Fig. 3 a). This approach avoided possible eukaryotic bridging proteins being present in the assay and was likely to identify direct interaction. The cloned (human) cDNAs were FLAG tagged and coexpressed with untagged human AID from the same plasmid. FLAG-PAF1 was co-isolated in AID immunoprecipitates, whereas CDC73 (Fig. 3 a), SSRP1 (not depicted), and LEO1 (not depicted) did not show robust association. The PAF1–AID association was specific (Fig. 3 a, lanes 4–6) and did not occur in the absence of AID-specific antisera (Fig. 3 a, lanes 7–9). A reciprocal IP experiment was also performed (unpublished data), verifying the AID–PAF1 association. To confirm the possible direct interaction between AID and PAF1, we performed classical pull-down analysis with recombinant AID and in vitro-produced PAF1. As shown in Fig. 3 b, PAF1 associated with AID but not APOBEC2, a member of the AID/APOBEC deaminase family. We also attempted to identify AID and SUPT5H association in the *Escherichia coli* and in vitro translation assays, but unlike the robust PAF1 association, were unable to demonstrate significant co-isolations (unpublished data).

To demonstrate that the AID–PAF1 association can provide a functional consequence in mammalian cells, we used a transcription reporter assay. PathDetect HeLa luciferase

reporter (HLR) cells harbor a luciferase transgene in their genome that can be activated by the PKA-phosphorylated CREB transcription factor. The presence of GAL4-binding sites (UAS) within the promoter allows for monitoring the effect of GAL4–fusion proteins on transcription. When GAL4 fusions of AID or AID mutant (E58Q) protein were transiently transfected, luciferase activity was enhanced nearly sixfold (Fig. 3 c, left). PAF1 and LEO1 chromatin IP (ChIP) of the transfected cells demonstrated that endogenous PAF1 and LEO1 were recruited to the locus upon AID expression. (Fig. 3 c, right), further underlying a more direct association between AID and the PAF complex.

Mapping the domain of AID that fostered this association was demonstrated by the use of AID–APOBEC2 chimeras, which substitute corresponding APOBEC2 peptide regions in place of AID peptide regions (Conticello et al., 2008). GFP-tagged AID, APOBEC2, or AID/APOBEC2 chimera proteins were coexpressed with Myc-peptide tagged human PAF1 in HEK293T cells and subjected to coIP. While IPs of AID and chimeras C and D showed co-purification of PAF1, APOBEC2 and chimera A and B failed to isolate PAF1 (Fig. 3 d), suggesting that the N-terminus of AID is responsible for the PAF1 association.

The PAF complex is required for functional CSR

Our finding that RNA pol II elongation factors associate with AID on chromatin, along with the previously established link of transcription being essential for SHM and CSR, provides an insight into the mechanism of AID activity at Ig loci. To determine the biological relevance of the PAF complex in CSR, we undertook knockdown experiments in murine B cells. CH12 cells were transduced with retrovirus-expressing shRNAs specific for the different subunits of the PAF complex. Transduced cells were stimulated in vitro, and their capacity to

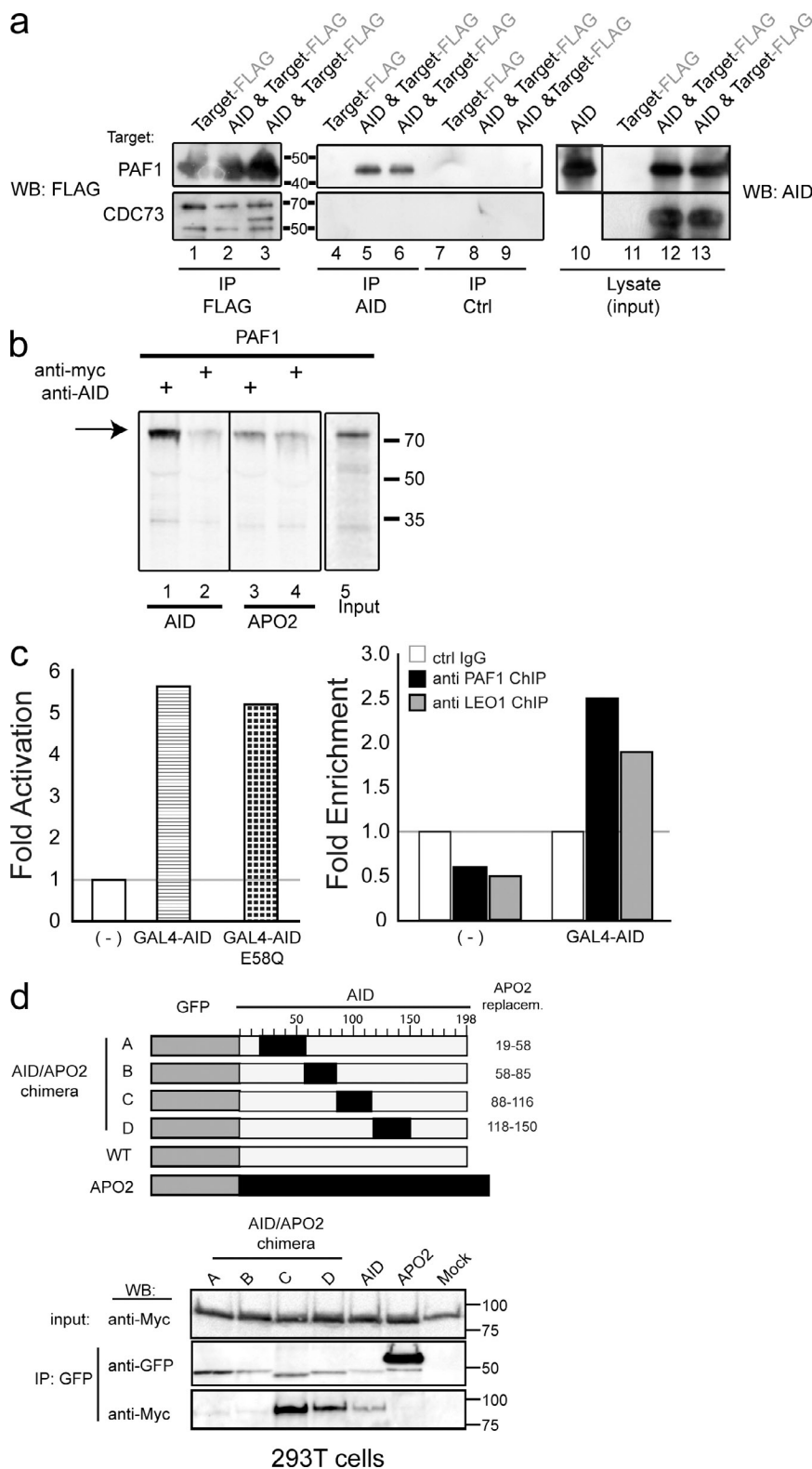


Figure 3. AID-PAF complex association via PAF1. (a) FLAG-PAF1 or FLAG-CDC73 were expressed alone (lanes 1, 4, 7, and 11) or from the same plasmid as AID (lanes 2, 3, 5, 6, 8, 9, 12, and 13) in *E. coli*. PAF1 is processed as a shortened (50 kD) fragment in bacteria. *E. coli* lysates were used for IP with anti-FLAG (lanes 1–3), anti-AID antibodies (lanes 4–6), or anti-MYC control antibody (lanes 7–9) and precipitates were analyzed in Western blots using anti-FLAG. AID expression in the lysates is shown in a separate blot on the right. Expressed protein and tag are indicated above lanes. (b) Pull-down assays were performed with ³⁵S-labeled, in vitro-translated PAF1 and recombinant AID (lanes 1 and 2) or APOBEC2 (APO2; lanes 3 and 4) purified from *E. coli*. Pull downs were performed using anti-AID (lanes 1 and 3) or control antibody anti-Myc (lane 2 and 4) and analyzed on SDS-PAGE, followed by autoradiography. Assays performed with APOBEC2 protein (lanes 3 and 4) served as controls. A 10% Input is shown in lane 5. (c, left) PathDetect HLR cells transfected with GAL4-AID and GAL4-AID E58Q were analyzed 30 h after transfection for transcription by luminometer (in triplicate). Data are representative of one of two independent experiments. (c, right) 30 h after transfection, chromatin was isolated and subjected to ChIP with anti-PAF1 or anti-LEO1 antibodies. Purified precipitated DNA was analyzed with qPCR (in triplicates), amplifying the 5' part of the luciferase gene. Data are representing one of two independent experiments. Mock transfected cells (-) served as references, and the control IgG was set to 1. (d, top) A scheme representing the AID/APOBEC2 chimera. The numbers indicate the amino acids of AID that are replaced in each chimera by corresponding APOBEC2 amino acids. (d, bottom) Myc-tagged PAF1 was coexpressed with GFP-AID, GFP-APOBEC2, or GFP-AID/APOBEC2 chimera A–D (Coticello et al., 2008) in HEK293T cells. GFP-proteins were immunopurified (anti-GFP) and IPs were probed with anti-MYC and anti-GFP antibodies (middle and bottom). For input, 5% of the lysate used for IP was analyzed for MYC-PAF1 expression (top).

undergo CSR to IgA was determined by flow cytometry (Fig. 4, a and b). As controls we used shRNAs specific for AID and SUPT5H, together with a nontarget shRNA control.

Knockdown efficiencies were determined by qRT-PCR (Fig. 4 c). Consistent with previous results (Pavri et al., 2010), we found that knockdown of AID and SUPT5H resulted in a significant reduction of CSR efficiency (Fig. 4, a and b). Knockdown of PAF1, LEO1, and CTR9 resulted in a similar reduction in the efficiency of CSR, which ranged from 31 to 35% (Fig. 4 b, gray bars), thus indicating the involvement of the PAF complex in CSR. No effects on viability, as determined by Topro-3 staining, were observed (not depicted). CDC73 depletion showed a reduction in CSR, but the change was not as significant as that of the other PAF

complex members. To verify the retrovirus shRNA knock-down effects on the PAF complex and possibly enhance the efficacy, we developed a lentivirus-based system. Although the overall switching efficiency was reduced even in the control samples, the lentiviral caused effect was much more pronounced, with a LEO1 knockdown reducing switching by >70% (Fig. 4 b and not depicted). This enhanced CSR inhibition by LEO1, can be explained, in part, by the more pronounced reduction of the target mRNA (Fig. 4c). Importantly, although the knockdown did not lead to a complete loss of the target, biological changes in CSR were observed.

As the PAF complex is part of the RNA pol II transcription machinery, the knockdown of its individual subunits could have broader influences on the cell than just altering AID's function at the IgH locus during CSR. We thus monitored the effect of knockdown on switch region transcription and AID expression. Although transcription at the donor switch region was not affected by the knockdown of any of the PAF complex subunits (Fig. 4 d), we found that knockdown of PAF1 and CTR9 resulted in altered levels of germline transcription at the acceptor switch region (Fig. 4 e). Furthermore, knockdown of PAF1, CTR9, CDC73, and SUPT5H resulted in a significant reduction in the level of AID mRNA (Fig. 4 f). Importantly, however, knockdown of LEO1 did not reduce AID mRNA expression (Fig. 4 f), nor reduce the levels of germline transcripts (Fig. 4, d and e), yet CSR was significantly reduced (Fig. 4, a and b); a finding that was confirmed with the lentivirus system. Because reduction in the expression of mismatch repair and base excision repair proteins, like UNG

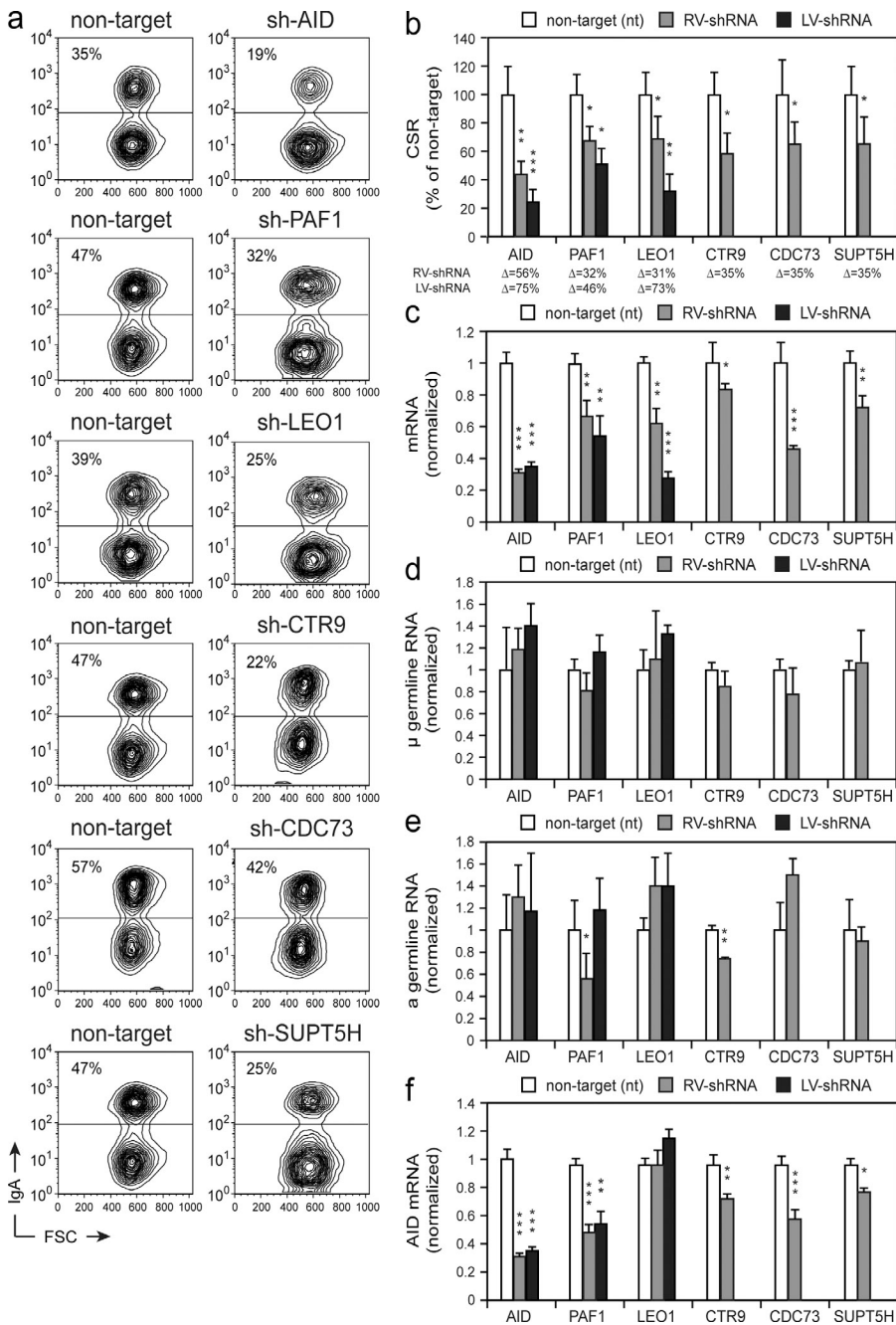


Figure 4. Knockdown of PAF complex members impairs CSR. (a) CH12 cells were retrovirally transduced with shRNAs for AID, PAF1, LEO1, CTR9, CDC73, SUPT5H, or a respective nontarget. After stimulation, IgA surface expression was monitored by flow cytometry, representative plots are shown. Numbers within the FACS plots indicate the percentage of IgA-positive cells. (b) The mean percentage (+ SD) of CSR in stimulated cells that were RV-shRNA transduced (gray bars) or LV-shRNA transduced (black bars), relative to the nontarget shRNA control (white bars) set to 100% from three independent experiments, is shown. The difference (Δ) in CSR efficiency between nontarget and target shRNA knockdown is shown below. Significance of a decrease versus nontarget is indicated by the following p-values: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$, based on Student's *t* test. Quantitative RT-PCR for AID, PAF1, LEO1, CTR9, CDC73, and SUPT5H transcripts (c), μ germline transcript (d), α germline transcript (e), and AID transcripts from cells transduced with RV-shRNA (gray bars) or LV-shRNA (black bars) shown relative to the corresponding nontarget shRNA control (f; white bars). Results are from three independent experiments. Transcript Ct values were normalized to CD79b or HPRT mRNA abundance and are presented relative to the levels in the nontarget shRNA negative control, set to 1. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$, based on Student's *t* test.

and MSH2/MSH6, could also explain the observed reduction in CSR, we monitored their expression level (by qRT-PCR) after knockdown of AID, PAF1, and LEO1. We were unable to identify any significant changes in mRNA levels (unpublished data).

Reducing the expression of the PAF complex proteins induced a loss in CSR, thereby identifying the PAF complex as a key component during Ig diversification. The observation that the core PAF protein LEO1 knockdown reduced CSR threefold, whereas not altering the expression of key transcript units, indicated that the PAF complex (or at least LEO1) plays a direct role in regulating AID function at the chromatin target.

PAF is present on the functional Ig allele of DT40 independently of AID

As a complex associated with active transcription, the PAF complex is present on numerous genes. To determine whether the PAF complex is recruited to an active Ig locus, we performed ChIP from DT40 chromatin using antibodies specific for PAF1 and LEO1 (Fig. 5). As in most B cells, in DT40 there is a strong allelic exclusion bias with only one of the two Ig light chain (lambda) alleles being active. By designing specific primers for the active (R, rearranged) and inactive (UR, unrearranged) allele (Fig. 5, schematic), we could identify PAF1 and LEO1 to be specifically located at the active allele. The PAF1 and LEO1 occupancy near the C domain (which is present on both alleles) was analogous to that of the previously described SUPT5H, and indicated a presence of the PAF complex outside of AID-targeted regions. This also led us to investigate if AID presence was necessary for PAF complex presence at the Ig locus, and we performed the same ChIP in AID-deficient DT40 cells (Harris et al., 2002). We found that PAF1 and LEO1 occupancy at the rearranged allele was not disrupted, and was even increased, by AID-deficiency (Fig. 5), indicating an AID-independent function

for the loading of the PAF complex proteins to Ig loci. We conclude that the PAF complex could serve as a binding platform for AID.

AID presence at S_{μ} is impaired by LEO1 knockdown

If the PAF complex can serve as a site for AID association at Ig loci, then reducing PAF expression should alter AID's occupancy at an Ig locus. To determine whether AID recruitment to the S_{μ} switch region is dependent on LEO1, ChIP experiments using an anti-AID antibody (Pavri et al., 2010) on unstimulated or stimulated transduced CH12 cells were performed (Jeevan-Raj et al., 2011). Before analysis, the cells had been transduced with a lentivirus expressing a shRNA specific for LEO1, AID, or a nontarget and were sorted for enhanced GFP expression. AID occupancy at the S_{μ} switch region was significantly reduced in LEO1 knockdown cells when compared with the nontarget control shRNA (Fig. 6). AID-ChIP signal was specific, as there was no significant difference in AID occupancy between unstimulated CH12 cells (not expressing AID) and stimulated cells expressing an shRNA specific for AID. We conclude that AID binding to S_{μ} is impaired by LEO1 knockdown. This result indicates that the functional mechanism of the PAF complex (at least LEO1) is to allow for AID to reside at an Ig locus during immune diversification.

DISCUSSION

Transcription has long been associated with AID-induced immune diversification. Early transgenic work demonstrated that the removal of the Ig promoter or enhancer elements abolished SHM (Betz et al., 1994). Mutation distribution across the V region of Ig genes indicated that AID-induced mutations are initiated 100–150 bp downstream of the transcription start site (TSS), and continue for ~1,500–2,000 bp (Rada et al., 2002). Recent work has identified a similar AID-induced mutation profile across non-Ig genes (Liu et al., 2008), although the extent and frequency of SHM on these non-Ig genes was much more restricted. This indicated that although transcription is

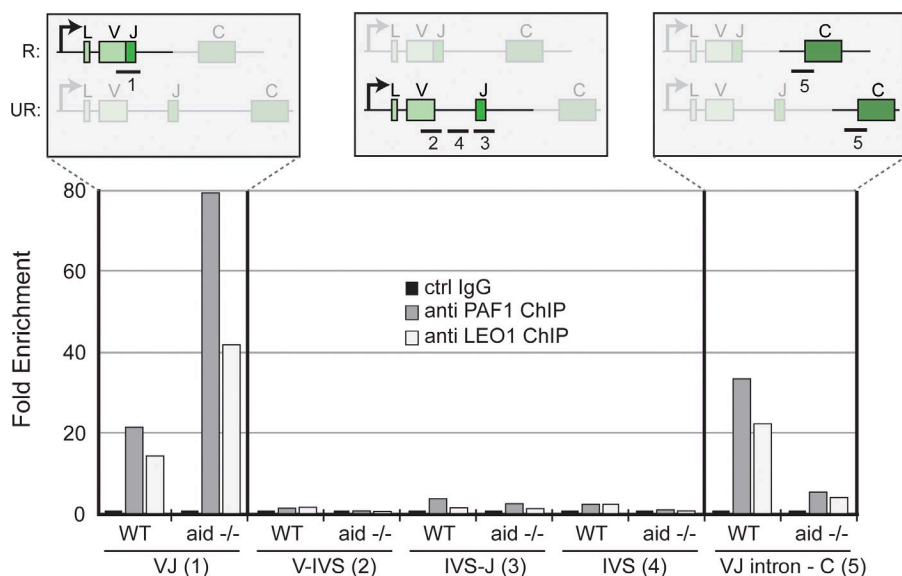


Figure 5. PAF complex presence at the active Ig allele independent of AID. DT40 ChIP was performed as described in the Materials and methods section. (top) Schematics of the R (rearranged) and UR (unrearranged) lambda alleles in DT40 cells. PCR amplifications are shown as lines: 1, rearranged V/J region allele; 2, 3, and 4, unrearranged V/J region allele; 5, both C region alleles. WT, parental DT40; aid $-/-$, AID KO has previously been described (Harris et al., 2002). ChIP was performed using anti-PAF1 (gray bars) and anti-LEO1 (white bars), and qPCR results (in triplicate) were compared with IgG control ChIP (black bars set to 1). One of two representative experiments is shown.

crucial, location and chromatin configuration also play a significant role, whereas sequence alone does not.

Several AID-associated proteins have been identified, some of which are linked directly to RNA processing (Conticello et al., 2008; Pavri et al., 2010; Stanlie et al., 2010; Basu et al., 2011; Okazaki et al., 2011), whereas others are important for subcellular localization (Patenaude et al., 2009; Maeda et al., 2010) or substrate accessibility (Chaudhuri et al., 2003). After fractionating B cells undergoing Ig diversification, we focused on the chromatin-bound AID and its physiological interactome (Fig. 1), which consisted of RNA pol II core (RNA pol II sub unit 1A and 2A) and associated proteins (SUPT5H), splicing factors (SF3A and 3B, Prp6, Prp4), RNA helicases, chromatin modifiers (SUPT6H, SSRP1 and SUPT16H), and an RNA pol II elongation complex (PAF complex; PAF1, LEO1, CTR9, CDC73). We verified these associations in DT40 and CH12F3 cells (Fig. 2, a and b), and demonstrated that PAF1 was the likely AID-interacting subunit within the PAF complex (Fig. 3). The biological significance of the AID-PAF complex association was shown by LEO1 knockdown in induced CH12 cells, where we observed reduced CSR without reducing AID or Ig transcript levels (Fig. 4). Mechanistically, at the Ig locus, the presence of the PAF complex (Fig. 5) enhanced AID occupancy (Fig. 6).

Transcription-coupled AID function

Genetically, transcription has been linked to SHM and CSR (Stavnezer-Nordgren and Sirlin, 1986; Peters and Storb,

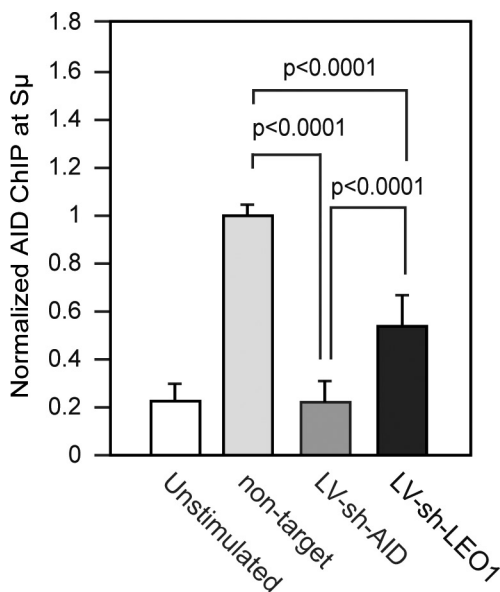


Figure 6. AID binding to S μ is impaired after LEO1 knockdown. Unstimulated and/or stimulated CH12 cells were transduced with a lentivirus expressing shRNAs specific for AID, LEO1, or a nontarget shRNA control. AID occupancy at the S μ switch region was ChIP analyzed using anti-AID antibodies. Normalized AID-ChIP data from three independent experiments assayed with two different primer sets is shown. For each sample, AID-ChIP values were normalized to the input control. AID-ChIP signal in cells expressing a nontarget shRNA control was set to 1. P-values, two-tailed Student's *t* test.

1996), whereas an AID RNA pol II association has subsequently been implicated (Nambu et al., 2003). During SHM and CSR, mutations do not occur until after promoter escape (>100 bp downstream of the start site), and because of this the processing of RNA is likely a mechanistic link to AID activity. This was confirmed by the discoveries of an association between the following: AID and CTNNB1, a protein of the splicing machinery, which occurs concomitantly during RNA pol II elongation (Conticello et al., 2008); AID and PTBP2, a splicing protein (Nowak et al., 2011); AID and SUPT5H, a protein known to associate with paused and elongating RNA pol II (Pavri et al., 2010); AID and SUPT4H, a factor known to associate with SUPT5H (Stanlie et al., 2012); AID and SUPT6H, a histone chaperone (Okazaki et al., 2011); CSR and SET1, a methyl-transferase for H3K4me3 (Stanlie et al., 2010); CSR and the FACT complex, a chromatin-modifying complex during RNA processing (Stanlie et al., 2010). Because of the involvement of the various RNA biogenesis and chromatin modification proteins in AID-induced Ig diversification, one cannot exclude the possibility that some of these factors serve multiple roles in directly controlling AID at the Ig locus, in changing the chromatin state of the Ig locus through the regulation of key factors, and in influencing the pathway and resolution of AID-lesions based on altered chromatin states.

The RNA pol II C-terminal domain (CTD) tail, which is temporally and spatially modified, serves as a platform for co-transcriptional mRNA maturation and chromatin modification. The PAF complex helps to set the right co-transcriptional chromatin marks, itself serving as a docking platform for the H2B ubiquitination machinery, as well as for setting H3K4me3 marks (Jaehning, 2010). H3K4me3 serves as an important mark in CSR (Wang et al., 2009; Stanlie et al., 2010), but is generally restricted to the 5' end of a gene, and replaced by H3K36me3 toward the 3' end of the gene. Both of these marks are induced upon transcriptional activation of S-regions (Wang et al., 2009), but at these loci, the H3K4me3 domain is extended, whereas onset of H3K36me3 is pushed back toward the 3' end. This correlates roughly with the cease of mutational load/AID activity in C regions (Wang et al., 2009). Our ChIP data in DT40 confirm that the machineries required to set the various marks are also skewed along the transcription unit during Ig diversification (Fig. 5). This data also confirms that occupancy by AID-associated factors does not equate to AID occupancy, given that the gross SUPT5H and RNA pol II occupancy profile is not altered for several hundred base pairs, extending into the C region (Pavri et al., 2010), and not all stalled genes are target for AID binding or mutation (Yamane et al., 2011). Furthermore, AID has been associated with TSS of non-Ig genes (Yamane et al., 2011), yet no functional relevance (e.g., AID-induced mutations) has been identified at these locations. Therefore, the current data of linking the early transcriptional events to AID association provides further insight into the establishment of 5' boundary-marks of

SHM, whereas the understanding of molecular mechanism for the 3' boundary remains less clear.

Overall, our work now provides the biochemical (and physiological) foundation for the aforementioned AID associations, while at the same time providing the molecular link (PAF complex) between early transcription elongation, marked by SUPT5H/SUPT4H, and downstream extended chromatin modifications dependent on FACT (SSRP1 and SUPT16H), SET1, and SUPT6H (Pavri et al., 2006; Fleming et al., 2008; Chen et al., 2009; Jaehning, 2010; Selth et al., 2010). A possible order of events at the Ig locus (Fig. 7) would entail the following: RNA pol II pausing after promoter escape and phosphorylation of its CTD tail; binding of the SUPT4H–SUPT5H complex to RNA pol II; recruitment of PAF complex to the holocomplex and initiation of histone modifications near the pause site (H2B mono-ubiquitination by the BRE1/RAD6 complex serves as a platform for SET1 complex for H3K4 trimethylation) and phosphorylation of CTD and

SUPT4H–SUPT5H complex by pTEFb; concomitant association of AID to the PAF–SUPT5H–RNA pol II complex, FACT complex recruitment and chromatin remodeling, SUPT6H association to the restarting polymerase; elongating/pausing transcription for enhanced AID resident time at Ig locus, RNA biogenesis, opening of chromatin and DNA for AID accessibility, recruitment of DNA repair factors to initiate SHM and CSR; hyperphosphorylation of the CTD, loss of AID association, and completion of RNA synthesis.

As mentioned above, several of the proposed proteins have been demonstrated to either associate with AID and/or play a role during Ig diversification. The identification of the nucleosome modifiers SUPT6H and FACT at the Ig locus, the demonstration that histone H3K4 trimethylation is necessary for CSR (Stanlie et al., 2010), and the correlation of H2Bser14 phosphorylation (Odegard and Schatz, 2006), H4K20 methylation (Schotta et al., 2008), H3 acetylation (Kuang et al., 2009; Wang et al., 2009), and H3K9 trimethylation (Chowdhury et al., 2008; Jeevan-Raj et al., 2011; Kuang et al., 2009; Wang et al., 2009) with Ig diversification indicates that the interplay of transcription and chromatin modification during AID-induced Ig diversification, although complex, is beginning to be unraveled. Although our data suggest that the predominant function of the PAF complex during SHM is to provide a site for AID association, we cannot exclude the possibility that reduced PAF activity also alters nucleosome marks needed for the resolution of AID-induced lesions, but more detailed future analysis may.

SHM versus CSR

The AID–PAF complex and AID–SUPT5H interactions were isolated from DT40 cells, which undergo SHM as well as gene conversion, but do not undergo CSR. Past work has implicated histone modification during SHM, but detailed understanding is still lacking, whereas H3K4me3 seems to play an important role during CSR (Stanlie et al., 2010). Our isolation of most of the required components for setting this mark during transcription would imply a similar requirement during V region diversification.

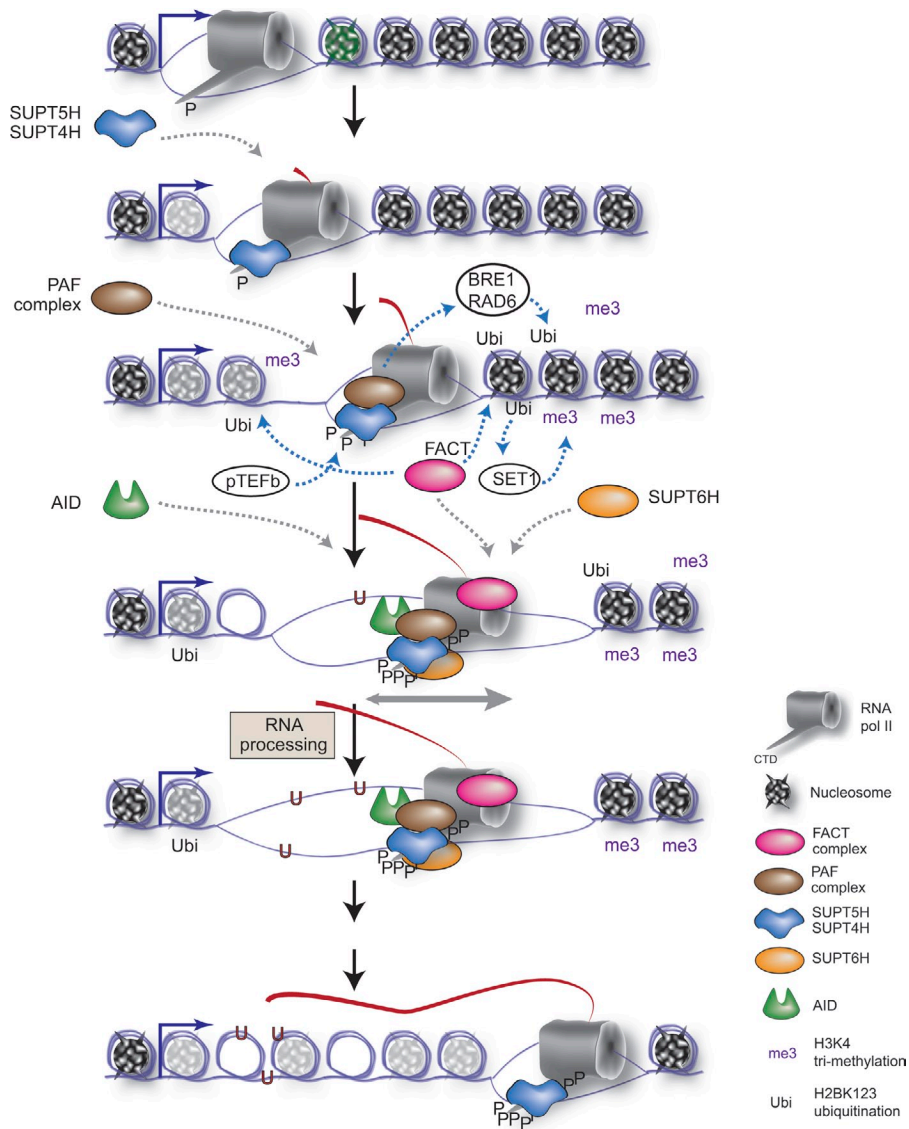


Figure 7. Model of the AID-linked transcriptional events at the Ig locus. Factors not currently identified to associate with AID or SHM/CSR are in white circles; other factors are indicated in the legend.

Furthermore, we also identified PAF interactions from cells undergoing CSR. On the other hand, there have been indications that SUPT6H (Okazaki et al., 2011), SUPT4H (Stanlie et al., 2012), and the FACT components (Stanlie et al., 2010) have different functionality during SHM and CSR, but detailed analysis from knock-outs and the endogenous SHM or CSR loci need to confirm the exact mechanisms.

Conclusion

Our work has provided biochemical and genetic insight into understanding the association of AID to the Ig locus. Our novel approach to isolate physiological AID-containing protein complexes only from chromatin has identified a new component, the PAF complex, as well as biochemically verified the significance of previously identified complexes (SUPT5H, SUPT6H, and FACT) in AID biology. Furthermore, our data extends the current model of AID gaining access to DNA by stalled RNA polymerase II to a more complex model, where AID is intimately and specifically linked with RNA pol II in the phase of pausing and elongation, surrounded by a specific chromatin environment defined by histone modification cascades.

The finding that AID interacts with the PAF and the RNA pol II elongation complexes is somewhat reminiscent of a model put forth by Peters and Storb (1996), where an unknown mutator (now known to be AID) would bind to initiating RNA pol II and travel along with the machinery during transcription elongation.

MATERIALS AND METHODS

Plasmids, cell lines, and antibodies. Plasmids were constructed using standard PCR and molecular biology techniques; sequences are available upon request. Tagging AID exon 5 in DT40 has been previously described (Pauklin et al., 2009), with the following modification: instead of a 3xFLAG-2xTEV-3xMyc tagged AID construct (AID-3FM), we also generated AID-3F (3xFLAG). Expression plasmids for GFP-AID, GFP-Apobec2, and GFP-AID/Apobec2 chimera A-D were obtained from the Neuberger Laboratory (Conticello et al., 2008). A CMV promoter-driven MYC-PAF1 expression vector was obtained by cloning the human PAF1 cDNA into pcDNATM3.1/myc-His (Invitrogen). For a complete list of antibodies used in this study please see [Table S3](#).

Chromatin AID-3FM and AID-3F isolation. Isolation was based on a previously described method (Aygün et al., 2008), with modifications. $1-2 \times 10^{10}$ DT40 cells (Pauklin et al., 2009) were collected by centrifugation at 1,200 rpm 4°C for 10 min, and cell pellets were washed twice with 50 ml cold 1xPBS. Cytoplasmic lysis: 5 times packed cell volume ($\sim 1 \mu\text{l}$ PCV = 10^6 cells) of Hypotonic Lysis Buffer (HLB; 10 mM Tris HCl [pH 7.5], 2 mM MgCl₂, 3 mM CaCl₂, and 0.32 M sucrose, protease inhibitor cocktail [Roche], and phosphatase inhibitor cocktail [Roche]) was added to the cell pellet, resuspended gently, and incubated for 12 min on ice. To the swollen cells, 10% Triton X-100 was added to a final concentration of 0.3%. The suspension was mixed and incubated for 3 min on ice, centrifuged for 5 min at 1,000 g at 4°C, and the supernatants (cytoplasmic fraction) were collected. Nuclear pellets were washed once in HLB + 0.3% Triton X-100, resuspended in 2xPCV LB-T (LB - 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 50 mM KCl, 2 mM MgCl₂, 1 mM EDTA, 10% Glycerol, protease inhibitor cocktail [Roche], phosphatase inhibitor cocktail [Roche], and 0.3% Triton X-100), and dounce homogenized with 30 strokes. The samples were incubated

with gentle agitation for 30 min at 4°C and ultracentrifuged at 33,000 g for 30 min at 4°C. The pellets were dounce homogenized until resistance was lost in 2xPCV LB-TB (LB-T + 150 U/ml Benzozase [VWR International]). The samples were incubated at room temperature for 30 min and ultracentrifuged as before. The supernatants (chromatin fraction) were subjected to a preclearing step with agarose beads before adding M2-affinity beads (Sigma-Aldrich) for IP. For PAF complex analysis, NaCl and KCl were doubled to give a final concentration of 300 mM, and the Triton X-100 concentration increased to 0.5%. For Western analysis of input, lysates between 0.5 and 3% of total lysate was loaded per lane.

Size exclusion chromatography. Chromatin extract of DT40 was prepared as described above. 1 ml of extract was loaded onto a Superdex 200 10/300 GL column, which had been equilibrated in LB and calibrated with standard proteins using Äkta Explorer (GE Healthcare). Fractions were collected at 1-ml volume steps using a 0.5 ml flow-rate, concentrated, and analyzed by Western blot.

FLAG-IP. Anti-FLAG M2 affinity beads (Sigma-Aldrich) were washed and equilibrated in LB-T. For chromatin fractions, 100 μl of M2 beads per 5×10^9 DT40 cells were incubated for 3-4 h with the chromatin at 4°C on a rotator and collected for 3 min at $300 \times g$ at 4°C. Beads were washed 5 times in 25 \times bead volume of LB-TF (LB-T supplemented with 0.5-1 $\mu\text{g}/\text{ml}$ 1xFLAG peptide N-DYDDDDK-C) and once with LB at 4°C for 10 min, followed by two elution steps in 4 \times bead volume of EB (LB + 500 $\mu\text{g}/\text{ml}$ 3xFLAG peptide N-MDYKDHGDYKDHDDIDYKDDDDK-C); first for 1 h at room temperature, and then over night at 4°C.

Mass spectrometry. Polyacrylamide gel slices (1-2 mm) containing IP-purified proteins were prepared for mass spectrometric analysis using the Janus liquid handling system (PerkinElmer). In brief, the excised protein gel pieces were placed in a well of a 96-well microtiter plate and de-stained with 50% vol/vol acetonitrile and 50 mM ammonium bicarbonate, and then reduced with 10 mM DTT and alkylated with 55 mM iodoacetamide. After alkylation, proteins were digested with 6 ng/ μL trypsin (Promega) overnight at 37°C. The resulting peptides were extracted in 2% vol/vol formic acid, 2% vol/vol acetonitrile. The digest was analyzed by nano-scale capillary LC-MS/MS using a nanoAcquity UPLC (Waters) to deliver a flow of ~ 300 nL/min. A C18 Symmetry 5 μm , 180 $\mu\text{m} \times 20$ mm μ -Precolumn (Waters), trapped the peptides before separation on a C18 BEH130 1.7 μm , 75 $\mu\text{m} \times 100$ mm analytical UPLC column (Waters). Peptides were eluted with a gradient of acetonitrile. The analytical column outlet was directly interfaced via a modified nano-flow electrospray ionization source, with a linear quadrupole ion trap mass spectrometer (LTQ XL/ETD, Thermo Fisher Scientific). LC-MS/MS information was collected using a data dependent analysis procedure. MS/MS scans were collected using an automatic gain control value of 4×10^4 and a threshold energy of 35 for collision induced dissociation. LC-MS/MS data were then searched against a protein database (UniProt Knowledge Base) using the Mascot search engine program (Matrix Science; Perkins et al., 1999). Database search parameters were set with a precursor tolerance of 1.0 D and a fragment ion mass tolerance of 0.8 D. One missed enzyme cleavage was allowed and variable modifications for oxidized methionine, carbamidomethyl cysteine, phosphorylated serine, threonine and tyrosine were included. MS/MS data were validated using the Scaffold program (Proteome Software, Inc.; Keller et al., 2002). All data were additionally interrogated manually.

Western blotting. The antibodies used are shown in [Table S3](#). Samples were prepared using standard procedures. Proteins were fractionated using NuPage Bis-Tris gels (Invitrogen) or homemade 10% PAA gels and transferred to Immobilon-P membranes (Millipore).

Nuclear extracts and coIP in murine B cells. Nuclear extracts were prepared using standard techniques from CH12F3 cells stably expressing AID^{FLAG-HA} (Jeevan-Raj et al., 2011). coIPs and Western blot analyses were

performed as previously described (Jeevan-Raj et al., 2011). For antibodies used, please see Table S3.

In vitro translation and coIP. AID-His-tagged protein was expressed in *E. coli* and purified as previously described (Coker et al., 2006). ³⁵S-labeled PAF1 was expressed using the TnT T7 Coupled Reticulocyte Lysate System (IVT) according to the manufacturer's instructions (Promega). Labeled protein mixture was mixed with 100 ng of AID or 300 ng of APOBEC2 protein for 1 h at room temperature and for 30 min at 4°C. Proteins were isolated by anti-AID (hAnp52-1; Conticello et al., 2008) or anti-Myc (9E10) coupled to Sepharose beads for 1 h at 4°C, washed 5 times in 1× TBS-T (50 mM Tris, pH 8.0, 300 mM NaCl, 1% Triton X-100, 2.5 mM TCEP, 2% BSA, and protease inhibitor [Roche]), resuspended in SDS-PAGE loading buffer, and separated on 12% Bis-Tris polyacrylamide gels (Invitrogen). Gels were dried, exposed, and analyzed using a Fuji Imaging system.

E. coli coIP. cDNAs of PAF1 and CDC73 were fused to a C-terminal FLAG tag in a pET DUET derivative coexpressing untagged human AID. Plasmids were transformed into BL21-CODONPLUS (DE3)-RIL cells (Stratagene), and protein expression was induced at 16°C with 1 mM IPTG in the presence of 0.1 mM ZnCl₂ (3 h). Cells were sonicated in TBS-T, debris were pelleted at 19,000 g, and IPs were performed using Sepharose-coupled anti-AID hAnp52-1 or anti-FLAG M2 antibodies. After five washes with TBS-T, the immunoprecipitates were analyzed by Western blot, using polyclonal anti-AID (Abcam) and monoclonal anti-FLAG (M2-HRP; Sigma-Aldrich) antibodies.

HeLa PathDetect analysis. The Stratagene PathDetect HLR Cell Line and GAL4-CREB and PKA expression vectors were purchased from Agilent. This HeLa-based Luciferase Reporter cell line contains a single locus with integrated synthetic minimal promoter and five yeast GAL4-binding sites (UAS) driving expression of the luciferase gene. Plasmids expressing GAL4-CREB, PKA, and GAL4-AID were transfected using Lipofectamine2000 (Invitrogen), and luciferase activity (in triplicate) was monitored 24–48 h after transfection according to the Luciferase Assay System manual (Promega). ChIP analysis using anti-PAF1, anti-LEO1, and a control IgG were done as follows: cells were cross-linked with 1% formaldehyde, nuclei were isolated and lysed in sonication buffer (1% SDS, 50 mM Tris HCl, pH 8.0, and 10 mM EDTA). After sonication in a BioRuptor, fragmented chromatin was diluted and incubated with antibodies or control IgGs overnight. Collected protein-DNA complexes were purified and analyzed by qPCR (in triplicate). ChIP data were normalized to the input signal for each chromatin sample, and control ChIPs were set to 1. For antibodies used, please see Table S3. For oligonucleotides used, please see Table S1. Two independent experiments were performed with one representative shown.

coIP in HEK293T cells. HEK293T cells were transfected with a plasmid expressing MYC-PAF1. 12 h after transfection cells were pooled to guarantee equal expression of MYC-PAF1, and split to allow a second transfection (12 h later) with expression plasmids for GFP-AID, GFP-APOBEC2, or GFP-AID/APO2 chimera. 24 h after second transfection, cells were lysed (lysis buffer: 50 mM Tris HCl, pH 8.0, 150 mM NaCl, 0.04% SDS, 1% NP-40), and GFP-protein expression in the lysates was estimated by scanning aliquots of a dilution series of the lysates with a Typhoon Scanner. Equal GFP and protein amounts were subjected to IP with anti-GFP at 4°C on, immunoprecipitates were collected with protein-A/G-Sepharose beads, and beads were washed and analyzed by Western blotting.

shRNA knockdown. Retroviral knockdown was done as follows: vectors containing shRNAs specific for SUPT5H, PAF1, LEO1, CTR9, CDC73, and the nontarget shRNA control were purchased from OriGene (Table S2). The hairpin sequence for AID (5'-ACCAGTCGCCATTATAATGCAA-3') was cloned into the LMP retroviral vector (Open Biosystems). CH12 cells were transduced as previously described (Barreto et al., 2003). Transduced cells were selected with 0.5 µg/ml puromycin for 1–5 d before induction

and sorting. Lentiviral knockdown was done as follows: The hairpin sequences for AID, PAF1, LEO1, and the nontarget shRNA control were cloned into the pLKO.1-puro-CMV-TurboGFP lentiviral vector (Sigma-Aldrich). Lenti-X 293T cells (Takara Bio Inc.) were transfected with vectors to produce the virus. 2 d later, CH12F3 cells were spin-infected with viral supernatants supplemented with 10 µg/ml polybrene (Santa Cruz Biotechnology). Cells were selected for 5 d with 1 µg/ml puromycin before induction. Hairpin sequences used are listed in Table S2.

Cell culture and flow cytometry. Retrovirally transduced CH12 cells were cultured with 5 ng/ml IL-4 (Sigma-Aldrich), 0.3 ng/ml TGF-β (R&D Systems), monoclonal 200 ng/ml anti-CD40 antibody (eBioscience), and 0.5 µg/ml puromycin and analyzed after 48–72 h for CSR (IgM to IgA) by flow cytometry, as previously described (Robert et al., 2009).

Real-time quantitative RT-PCR. RNA and cDNA were prepared using standard techniques. qPCR was performed in triplicates using SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich) and a LightCycler 480 (Roche). Transcript quantities were calculated relative to standard curves and normalized to CD79b or HPRT mRNA. For primers see Table S1.

ChIP from DT40 and CH12 cells. In brief, DT40 cells were treated and analyzed as for the ChIP in the HeLa PathDetect analysis section. For antibodies used please see Table S3. For oligonucleotides used please see Table S1. Two independent experiments were performed, with one representative shown. For quantitative AID-ChIP from shRNA knockdowns: CH12 cells were transduced with a lentivirus expressing shRNAs specific for AID, LEO1 and a nontarget control. Cells were stimulated for 48 h and sorted for enhanced GFP expression using a FACS Aria II (BD) and/or FACSVantage SE (BD) cell sorters before ChIP analysis. Cells were cross-linked with 1% formaldehyde for 10 min. Chromatin was prepared and immunoprecipitated with an anti-AID antibody (Pavri et al., 2010) and analyzed by quantitative PCR as previously described (Jeevan-Raj et al., 2011). Raw data were normalized to the input signal for each sample. AID-ChIP signal in cells expressing a nontarget shRNA control was assigned an arbitrary value of 1. Statistical analysis was performed using a two-tailed Student's *t* test.

Online supplemental material. Fig. S1 is a schematic of how the isolation and analysis of the AID-associated complex was undertaken and a table of peptide IDs. Fig. S2 shows the AID interactome. Table S1 shows primer sequences. Table S2 lists shRNA sequences. Table S3 lists antibodies used in this study. Online supplemental material is available at <http://www.jem.org/cgi/content/full/jem.20112145/DC1>.

We would like to thank the members of the Petersen-Mahrt and Reina-San-Martin laboratories for discussions; Cancer Research UK (CRUK) cell services for performing DT40 growth; Michel Nussenzweig for the anti-AID antibody and Anna Gazumyan for advice on AID-ChIP; Gudrun Bachmann and Dafne Solera for help in the generation of cell lines; Claudine Ebel for cell sorting; and Jesper Svejstrup for discussion and critical reading of the manuscript.

K.L. Willman, S. Pauklin, G. Rangam, M.T. Simon, S. Maslen, and M. Skehel were supported by CRUK. S. Pauklin was supported in part by SA Archimedes-Estonian Foundation of European Union Education and Research. K.-M. Schmitz is supported by a Marie Curie FP 7 fellowship. S. Milosevic was supported by la Ligue Contre le Cancer, France. B. Reina-San-Martin is an AVENIR-INSERM young investigator. This work was supported by grants to B. Reina-San-Martin from the Agence Nationale pour la Recherche (ANR-07-MIME-004-01) and the Institut National de la Santé et de la Recherche Médicale (INSERM), and to S.K. Petersen-Mahrt from Istituto FIRC di Oncologia Molecolare, Italy (IFOM) and CRUK.

The authors have no conflicting financial interests.

Author contributions: K.L. Willman, S. Milosevic, S. Pauklin, K.-M. Schmitz, G. Rangam, M.T. Simon, I. Robert, V. Heyer, and E. Schiavo performed experiments. S. Pauklin performed the initial AID on chromatin fractionation and isolation. S. Maslen and M. Skehel performed mass spectrometry analysis. K.L. Willman,

S. Milosevic, K.-M. Schmitz, B. Reina-San-Martin, and SKPM analyzed the data. K.-M. Schmitz, K.L. Willman, S. Milosevic, B. Reina-San-Martin, and S.K. Petersen-Mahrt wrote the paper. K.L. Willman, SM, S. Pauklin, K.-M. Schmitz, B. Reina-San-Martin, and S.K. Petersen-Mahrt designed the experiments. S.K. Petersen-Mahrt conceived the approach.

Submitted: 10 October 2011

Accepted: 22 August 2012

REFERENCES

- Aygun, O., J. Svejstrup, and Y. Liu. 2008. A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin. *Proc. Natl. Acad. Sci. USA.* 105:8580–8584. <http://dx.doi.org/10.1073/pnas.0804424105>
- Barreto, V., B. Reina-San-Martin, A.R. Ramiro, K.M. McBride, and M.C. Nussenzweig. 2003. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol. Cell.* 12:501–508. [http://dx.doi.org/10.1016/S1097-2765\(03\)00309-5](http://dx.doi.org/10.1016/S1097-2765(03)00309-5)
- Basu, U., F.L. Meng, C. Keim, V. Grinstein, E. Pefanis, J. Eccleston, T. Zhang, D. Myers, C.R. Wasserman, D.R. Wesemann, et al. 2011. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell.* 144:353–363. <http://dx.doi.org/10.1016/j.cell.2011.01.001>
- Betz, A.G., C. Milstein, A. González-Fernández, R. Pannell, T. Larson, and M.S. Neuberger. 1994. Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. *Cell.* 77:239–248. [http://dx.doi.org/10.1016/0092-8674\(94\)90316-6](http://dx.doi.org/10.1016/0092-8674(94)90316-6)
- Bransteitter, R., P. Pham, M.D. Scharff, and M.F. Goodman. 2003. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. USA.* 100:4102–4107. <http://dx.doi.org/10.1073/pnas.0730835100>
- Brar, S.S., M. Watson, and M. Diaz. 2004. Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. *J. Biol. Chem.* 279:26395–26401. <http://dx.doi.org/10.1074/jbc.M403503200>
- Chaudhuri, J., M. Tian, C. Khuong, K. Chua, E. Pinaud, and F.W. Alt. 2003. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature.* 422:726–730. <http://dx.doi.org/10.1038/nature01574>
- Chen, Y., Y. Yamaguchi, Y. Tsugen, J. Yamamoto, T. Yamada, M. Nakamura, K. Hisatake, and H. Handa. 2009. DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation. *Genes Dev.* 23:2765–2777. <http://dx.doi.org/10.1101/gad.1834709>
- Chowdhury, M., O. Forouhi, S. Dayal, N. McCloskey, H.J. Gould, G. Felsenfeld, and D.J. Fear. 2008. Analysis of intergenic transcription and histone modification across the human immunoglobulin heavy-chain locus. *Proc. Natl. Acad. Sci. USA.* 105:15872–15877. <http://dx.doi.org/10.1073/pnas.0808462105>
- Coker, H.A., H.D. Morgan, and S.K. Petersen-Mahrt. 2006. Genetic and in vitro assays of DNA deamination. *Methods Enzymol.* 408:156–170. [http://dx.doi.org/10.1016/S0076-6879\(06\)08010-4](http://dx.doi.org/10.1016/S0076-6879(06)08010-4)
- Conticello, S.G., K. Ganesh, K. Xue, M. Lu, C. Rada, and M.S. Neuberger. 2008. Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNB1. *Mol. Cell.* 31:474–484. <http://dx.doi.org/10.1016/j.molcel.2008.07.009>
- Di Noia, J.M., and M.S. Neuberger. 2007. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* 76:1–22. <http://dx.doi.org/10.1146/annurev.biochem.76.061705.090740>
- Fleming, A.B., C.-F. Kao, C. Hillyer, M. Pikaart, and M.A. Osley. 2008. H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation. *Mol. Cell.* 31:57–66. <http://dx.doi.org/10.1016/j.molcel.2008.04.025>
- Ganesh, K., S. Adam, B. Taylor, P. Simpson, C. Rada, and M. Neuberger. 2011. CTNNB1 is a novel nuclear localization sequence-binding protein that recognizes RNA-splicing factors CDC5L and Prp31. *J. Biol. Chem.* 286:17091–17102. <http://dx.doi.org/10.1074/jbc.M110.208769>
- Harris, R.S., J.E. Sale, S.K. Petersen-Mahrt, and M.S. Neuberger. 2002. AID is essential for immunoglobulin V gene conversion in a cultured B cell line. *Curr. Biol.* 12:435–438. [http://dx.doi.org/10.1016/S0960-9822\(02\)00717-0](http://dx.doi.org/10.1016/S0960-9822(02)00717-0)
- Ito, S., H. Nagaoka, R. Shinkura, N. Begum, M. Muramatsu, M. Nakata, and T. Honjo. 2004. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl. Acad. Sci. USA.* 101:1975–1980. <http://dx.doi.org/10.1073/pnas.0307335101>
- Jaehning, J.A. 2010. The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochim. Biophys. Acta.* 1799:379–388. <http://dx.doi.org/10.1016/j.bbagr.2010.01.001>
- Jeevan-Raj, B.P., I. Robert, V. Heyer, A. Page, J.H. Wang, F. Cammas, F.W. Alt, R. Losson, and B. Reina-San-Martin. 2011. Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *J. Exp. Med.* 208:1649–1660. <http://dx.doi.org/10.1084/jem.20110118>
- Keller, A., A.I. Nesvizhskii, E. Kolker, and R. Aebersold. 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* 74:5383–5392. <http://dx.doi.org/10.1021/ac025747h>
- Kim, J., M. Guermah, and R.G. Roeder. 2010. The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. *Cell.* 140:491–503. <http://dx.doi.org/10.1016/j.cell.2009.12.050>
- Kuang, F.L., Z. Luo, and M.D. Scharff. 2009. H3 trimethyl K9 and H3 acetyl K9 chromatin modifications are associated with class switch recombination. *Proc. Natl. Acad. Sci. USA.* 106:5288–5293. <http://dx.doi.org/10.1073/pnas.0901368106>
- Liu, M., J.L. Duke, D.J. Richter, C.G. Vinuesa, C.C. Goodnow, S.H. Kleinstein, and D.G. Schatz. 2008. Two levels of protection for the B cell genome during somatic hypermutation. *Nature.* 451:841–845. <http://dx.doi.org/10.1038/nature06547>
- Maeda, K., S.K. Singh, K. Eda, M. Kitabatake, P. Pham, M.F. Goodman, and N. Sakaguchi. 2010. GANP-mediated recruitment of activation-induced cytidine deaminase to cell nuclei and to immunoglobulin variable region DNA. *J. Biol. Chem.* 285:23945–23953. <http://dx.doi.org/10.1074/jbc.M110.131441>
- McBride, K.M., V. Barreto, A.R. Ramiro, P. Stavropoulos, and M.C. Nussenzweig. 2004. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J. Exp. Med.* 199:1235–1244. <http://dx.doi.org/10.1084/jem.20040373>
- Nambu, Y., M. Sugai, H. Gonda, C.-G. Lee, T. Katakai, Y. Agata, Y. Yokota, and A. Shimizu. 2003. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science.* 302:2137–2140. <http://dx.doi.org/10.1126/science.1092481>
- Nowak, U., A.J. Matthews, S. Zheng, and J. Chaudhuri. 2011. The splicing regulator PTBP2 interacts with the cytidine deaminase AID and promotes binding of AID to switch-region DNA. *Nat. Immunol.* 12:160–166. <http://dx.doi.org/10.1038/ni.1977>
- Odegard, V.H., and D.G. Schatz. 2006. Targeting of somatic hypermutation. *Nat. Rev. Immunol.* 6:573–583. <http://dx.doi.org/10.1038/nri1896>
- Okazaki, I.-M., K. Okawa, M. Kobayashi, K. Yoshikawa, S. Kawamoto, H. Nagaoka, R. Shinkura, Y. Kitawaki, H. Taniguchi, T. Natsume, et al. 2011. Histone chaperone Spt6 is required for class switch recombination but not somatic hypermutation. *Proc. Natl. Acad. Sci. USA.* 108:7920–7925. <http://dx.doi.org/10.1073/pnas.1104423108>
- Patenaude, A.M., A. Orthwein, Y. Hu, V.A. Campo, B. Kavli, A. Buschiazio, and J.M. Di Noia. 2009. Active nuclear import and cytoplasmic retention of activation-induced deaminase. *Nat. Struct. Mol. Biol.* 16:517–527. <http://dx.doi.org/10.1038/nsmb.1598>
- Pauklin, S., I.V. Sernández, G. Bachmann, A.R. Ramiro, and S.K. Petersen-Mahrt. 2009. Estrogen directly activates AID transcription and function. *J. Exp. Med.* 206:99–111. <http://dx.doi.org/10.1084/jem.20080521>
- Pavri, R., B. Zhu, G. Li, P. Trojer, S. Mandal, A. Shilatfard, and D. Reinberg. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell.* 125:703–717. <http://dx.doi.org/10.1016/j.cell.2006.04.029>

- Pavri, R., A. Gazumyan, M. Jankovic, M. Di Virgilio, I. Klein, C. Ansarah-Sobrinho, W. Resch, A. Yamane, B. Reina San-Martin, V. Barreto, et al. 2010. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell*. 143:122–133. <http://dx.doi.org/10.1016/j.cell.2010.09.017>
- Perkins, D.N., D.J. Pappin, D.M. Creasy, and J.S. Cottrell. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 20:3551–3567.
- Peters, A., and U. Storb. 1996. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity*. 4:57–65. [http://dx.doi.org/10.1016/S1074-7613\(00\)80298-8](http://dx.doi.org/10.1016/S1074-7613(00)80298-8)
- Petersen-Mahrt, S. 2005. DNA deamination in immunity. *Immunol. Rev.* 203:80–97. <http://dx.doi.org/10.1111/j.0105-2896.2005.00232.x>
- Petersen-Mahrt, S.K., R.S. Harris, and M.S. Neuberger. 2002. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature*. 418:99–103. <http://dx.doi.org/10.1038/nature00862>
- Rada, C., J.M. Jarvis, and C. Milstein. 2002. AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc. Natl. Acad. Sci. USA*. 99:7003–7008. <http://dx.doi.org/10.1073/pnas.092160999>
- Rajewsky, K. 1996. Clonal selection and learning in the antibody system. *Nature*. 381:751–758. <http://dx.doi.org/10.1038/381751a0>
- Robert, I., F. Dantzer, and B. Reina-San-Martin. 2009. Parp1 facilitates alternative NHEJ, whereas Parp2 suppresses IgH/c-myc translocations during immunoglobulin class switch recombination. *J. Exp. Med.* 206:1047–1056. <http://dx.doi.org/10.1084/jem.20082468>
- Schotta, G., R. Sengupta, S. Kubicek, S. Malin, M. Kauer, E. Callén, A. Celeste, M. Pagani, S. Opravil, I.A. De La Rosa-Velazquez, et al. 2008. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev.* 22:2048–2061. <http://dx.doi.org/10.1101/gad.476008>
- Selth, L.A., S. Sigurdsson, and J.Q. Svejstrup. 2010. Transcript Elongation by RNA Polymerase II. *Annu Rev Biochem*.
- Stanlie, A., M. Aida, M. Muramatsu, T. Honjo, and N.A. Begum. 2010. Histone3 lysine4 trimethylation regulated by the facilitates chromatin transcription complex is critical for DNA cleavage in class switch recombination. *Proc. Natl. Acad. Sci. USA*. 107:22190–22195. <http://dx.doi.org/10.1073/pnas.1016923108>
- Stanlie, A., N.A. Begum, H. Akiyama, and T. Honjo. 2012. The DSIF subunits Spt4 and Spt5 have distinct roles at various phases of immunoglobulin class switch recombination. *PLoS Genet.* 8:e1002675. <http://dx.doi.org/10.1371/journal.pgen.1002675>
- Stavnezer, J., J.E. Guikema, and C.E. Schrader. 2008. Mechanism and regulation of class switch recombination. *Annu. Rev. Immunol.* 26:261–292. <http://dx.doi.org/10.1146/annurev.immunol.26.021607.090248>
- Stavnezer-Nordgren, J., and S. Sirlin. 1986. Specificity of immunoglobulin heavy chain switch correlates with activity of germline heavy chain genes prior to switching. *EMBO J.* 5:95–102.
- Wang, L., R. Wuerffel, S. Feldman, A.A. Khamlichi, and A.L. Kenter. 2009. S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. *J. Exp. Med.* 206:1817–1830. <http://dx.doi.org/10.1084/jem.20081678>
- Yamane, A., W. Resch, N. Kuo, S. Kuchen, Z. Li, H.-W. Sun, D.F. Robbani, K. McBride, M.C. Nussenzweig, and R. Casellas. 2011. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat. Immunol.* 12:62–69. <http://dx.doi.org/10.1038/ni.1964>

The cohesin complex regulates immunoglobulin class switch recombination

Anne-Sophie Thomas-Claudepierre, Ebe Schiavo, Vincent Heyer, Marjorie Fournier, Adeline Page, Isabelle Robert, and Bernardo Reina-San-Martin

Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de la Santé et de la Recherche Médicale (INSERM) U964/Centre National de la Recherche Scientifique (CNRS) UMR 7104/Université de Strasbourg, 67404 Illkirch, France

Immunoglobulin (Ig) class switch recombination (CSR) is initiated by the transcription-coupled recruitment of activation-induced cytidine deaminase (AID) to switch regions and by the subsequent generation of double-stranded DNA breaks (DSBs). These DNA breaks are ultimately resolved through the nonhomologous end joining (NHEJ) pathway. We show that during CSR, AID associates with subunits of cohesin, a complex previously implicated in sister chromatid cohesion, DNA repair, and the formation of DNA loops between enhancers and promoters. Furthermore, we implicate the cohesin complex in the mechanism of CSR by showing that cohesin is dynamically recruited to the $S\mu$ - $C\mu$ region of the IgH locus during CSR and that knockdown of cohesin or its regulatory subunits results in impaired CSR and increased usage of microhomology-based end joining.

CORRESPONDENCE

Bernardo Reina-San-Martin:
reinab@igbmc.fr

Abbreviations used: 3'RR, 3' regulatory region; AID, activation-induced cytidine deaminase; CSR, class switch recombination; DSB, double-stranded DNA break; IgH, Ig heavy chain; MudPIT, Multidimensional Protein Identification Technology; NHEJ, nonhomologous end joining; SHM, somatic hypermutation.

During immune responses, B cells diversify their receptors through somatic hypermutation (SHM) and class switch recombination (CSR). SHM introduces mutations in Ig variable regions that modify the affinity of the receptor for its cognate antigen (Di Noia and Neuberger, 2007). CSR replaces the antibody isotype expressed (from IgM to IgG, IgE, or IgA), providing novel antibody effector functions (Chaudhuri et al., 2007). Mechanistically, SHM and CSR are initiated by activation-induced cytidine deaminase (AID), an enzyme which deaminates cytosines in both strands of transcribed DNA substrates (Petersen-Mahrt et al., 2002; Basu et al., 2011). AID-induced DNA deamination is then processed to trigger mutations in variable regions during SHM or to generate double-stranded DNA break (DSB) intermediates in switch (S) regions during CSR (Chaudhuri et al., 2007; Di Noia and Neuberger, 2007). These breaks activate the DNA damage response (Ramiro et al., 2007) and are resolved through classical and alternative nonhomologous end joining (NHEJ; Stavnez et al., 2010).

CSR is a transcription-dependent, long-range recombination that occurs at the Ig heavy chain (IgH) locus and that involves the joining

of two S regions, which may be separated by several hundreds of kilobase pairs. For CSR to succeed, donor and acceptor S regions must be brought into close proximity. This is believed to occur through three-dimensional conformational changes involving the generation of transcription-coupled DNA loops (Kenter et al., 2012). Nevertheless, the precise mechanisms controlling these conformational changes remain to be elucidated.

The cohesin complex has been described to play a prominent role in sister chromatid cohesion during cell division, in favoring DNA repair by homologous recombination (Nasmyth and Haering, 2009), in modulating gene expression (Dorsett, 2009), and in promoting the transcription-coupled formation of long-range DNA loop structures (Kagey et al., 2010). In addition, cohesin and the transcriptional insulator CTCF (Dorsett, 2009; Nasmyth and Haering, 2009) have been shown to control the RAG1/2-dependent rearrangement of antigen receptor genes during early B and T lymphocyte development by mechanisms involving the regulation of transcription and formation of long-range in

A.-S. Thomas-Claudepierre and E. Schiavo contributed equally to this paper.

© 2013 Thomas-Claudepierre et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

cis DNA interactions (Degner et al., 2011; Guo et al., 2011; Seitan et al., 2011). Here, we have examined the role of cohesin in mature B cells undergoing CSR.

RESULTS AND DISCUSSION

Nuclear and chromatin-bound AID associate with cohesin

We have previously shown that nuclear AID exists in a large molecular weight complex containing proteins that are required for CSR (Jeevan-Raj et al., 2011). To further characterize this complex and investigate the functional role of novel AID partners in CSR, we have performed additional coimmunoprecipitation experiments coupled to identification by mass spectrometry. Nuclear and chromatin extracts prepared from CH12 cells expressing a full-length N-terminally tagged AID protein (AID^{Flag-HA}) or the epitope tags alone (Flag-HA) as negative controls were immunoprecipitated using an anti-Flag antibody. Eluted proteins were submitted for identification by mass spectrometry. Among the proteins identified, we found multiple AID partners previously implicated in CSR and/or SHM (Table S1). In addition, we found several proteins with no known function in CSR (Table S2), including subunits of the cohesin, condensin, Smc5/6 complex and Ino80 complexes. Given the described role for cohesin in mediating long-range recombination during B cell and T cell differentiation, we focused on the potential role of cohesin in CSR. The association between AID and the cohesin complex subunits (Smc1, Smc3, Nipbl, and Wapal) was confirmed by reciprocal coimmunoprecipitations and Western blotting in the nuclear (Fig. 1 A) and chromatin (Fig. 1 B) fractions and was specific, as they did not coprecipitate with an irrelevant tagged protein (EGFP^{Flag-HA}; Fig. 1 C). Importantly, these interactions were not mediated by nonspecific nucleic acid binding, as extracts and immunoprecipitations were done in the presence of the benzonase nuclease. We conclude that endogenous subunits of the cohesin complex associate with a fraction of nuclear and chromatin-bound tagged AID through interactions that do not involve nonspecific nucleic acid binding.

Smc1 and Smc3 are dynamically recruited to the IgH locus during CSR

To determine whether cohesin is recruited to the IgH locus in B cells undergoing CSR, we performed ChIP-Seq experiments

on chromatin prepared from resting or activated splenic B cells isolated from wild-type mice and using antibodies specific for Smc1, Smc3, and CTCF (Fig. 2). In resting B cells, we found that Smc1, Smc3, and CTCF are co-recruited to the 3' regulatory region (3'RR; Fig. 2 A). This is consistent with published ChIP data on CTCF (Chatterjee et al., 2011) in mature B cells and ChIP-Seq results for CTCF and cohesin (Rad21) in Rag1-deficient pro-B cells (Degner et al., 2011). A sharp peak of CTCF, Smc1, and Smc3 binding was observed at C α . This peak occurred over a region containing a predicted DNaseI hypersensitive site and a CTCF consensus motif (Nakahashi et al., 2013). No significant enrichment was observed at the E μ enhancer S μ or S γ 1 (Fig. 2 A). After stimulation, under conditions that induce CSR to IgG1, we found that Smc1 and Smc3 are significantly co-recruited, independently of CTCF, to a region spanning from the 5' end of the donor switch region (S μ) to the 3' end of the C μ constant region that did not comprise the E μ enhancer (Fig. 2 B). Surprisingly, we failed to detect a reproducible recruitment of Smc1 or Smc3 over the S γ 1 switch region (Fig. 2 B), suggesting that Smc1 and Smc3 are not recruited to the acceptor switch region upon activation. It is possible, however, that our cell culture conditions (in which ~15–20% of the cells switch to IgG1) are not robust enough to detect a specific enrichment. Consistent with this, we were unable to reproducibly detect a specific enrichment of AID at S γ 1 by ChIP-qPCR (Fig. 2 E).

The ChIP-Seq signal obtained in resting and activated B cells for Smc1 and Smc3 (Fig. 2, A and B) is consistent with the fact that they are known to exist as a heterodimer and was reproducible and specific, as we did not observe any significant enrichment at the IgH locus when using an IgG antibody as a negative control (Fig. 2, A and B). The recruitment of Smc1 and Smc3 at the IgH locus only partially correlated with that reported for AID (Yamane et al., 2011) and is consistent with the fact that only a fraction of chromatin-bound AID associates with the cohesin complex (Fig. 1 B). This suggests that cohesin is not a targeting factor for AID. The recruitment of Smc1, Smc3, and CTCF in resting and activated B cells observed by ChIP-Seq (Fig. 2, A and B) was confirmed by additional independent analytical-scale ChIP-qPCR experiments, using primer pairs at individual locations across

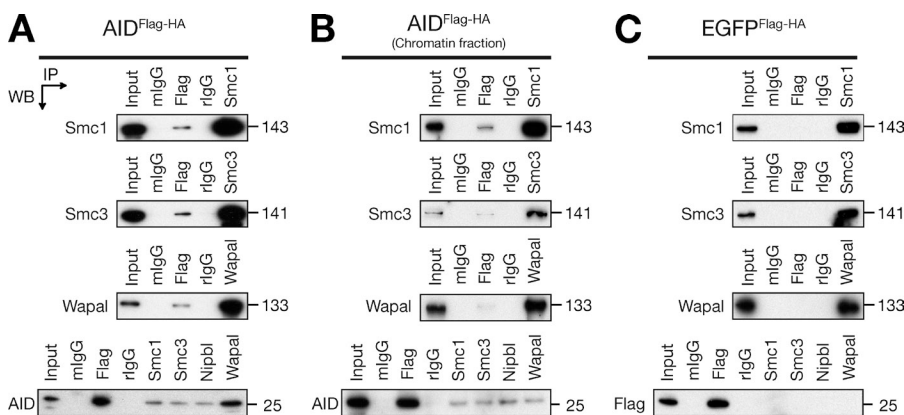


Figure 1. Nuclear AID associates with cohesin subunits. Nuclear extracts (A and C) and chromatin fractions (B) prepared from CH12 cells expressing AID^{Flag-HA} (A and B) or EGFP^{Flag-HA} (C) were immunoprecipitated and blotted with antibodies specific for Flag, AID, Smc1, Smc3, Wapal, and Nipbl. Note that the Nipbl antibody works only on immunoprecipitation. Input represents 1% of material used. Theoretical molecular masses in kilodaltons are indicated. Data are representative of three independent experiments.

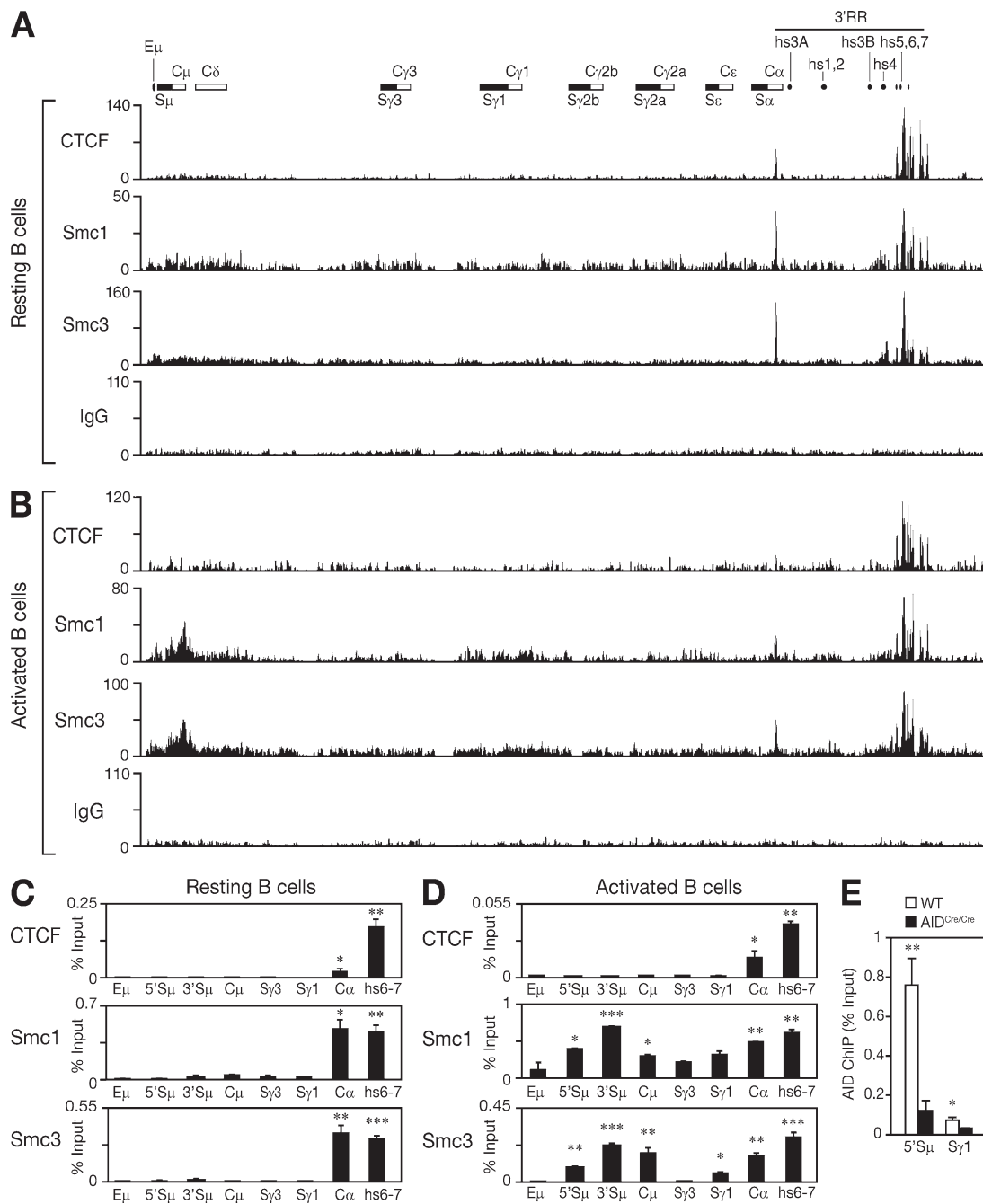


Figure 2. Smc1 and Smc3 are dynamically recruited to the IgH locus during CSR. UCSC genome browser screenshots showing the ChIP-Seq binding profiles of CTCF, Smc1, Smc3, and IgG (negative control) at the IgH locus (chr12:114,438,857–114,669,149) in resting (A) and activated (B; with LPS + IL-4) B cells isolated from wild-type mice. A schematic map of the IgH locus indicates the switch regions (black boxes), the constant region exons (white boxes), the E μ enhancer, and the DNaseI hypersensitive sites (hs) located in the 3'RR. Similar ChIP-Seq profiles were observed in an additional biological replicate experiment for Smc3 that was conducted in resting and activated B cells (not depicted). ChIP-Seq results were verified by analytical-scale ChIP-qPCR experiments performed on chromatin prepared from 10⁷ splenic resting (C) and activated (D) B cells. qPCR was performed at several locations across the IgH locus using primers listed in Table S4. Results are expressed as percent input and are representative of two independent biological replicate experiments. Mean of triplicate samples (+SD) is shown. Statistical significance versus S γ 3 (two-tailed Student's *t* test) is indicated: *, *P* ≤ 0.05; **, *P* ≤ 0.01; ***, *P* ≤ 0.001. Additional statistical analyses across the locus and between resting and activated B cells are shown in Table S5. (E) ChIP analysis for AID occupancy at the S μ and S γ 1 switch regions in wild-type and AID^{Cre/Cre} B cells cultured in vitro with LPS + IL-4 for 60 h. Results are expressed as percent input. Mean of triplicate samples (+SD) is shown. Statistical significance versus AID^{Cre/Cre} was determined by a two-tailed Student's *t* test. *, *P* ≤ 0.05; **, *P* ≤ 0.01. Results are representative of four independent experiments.

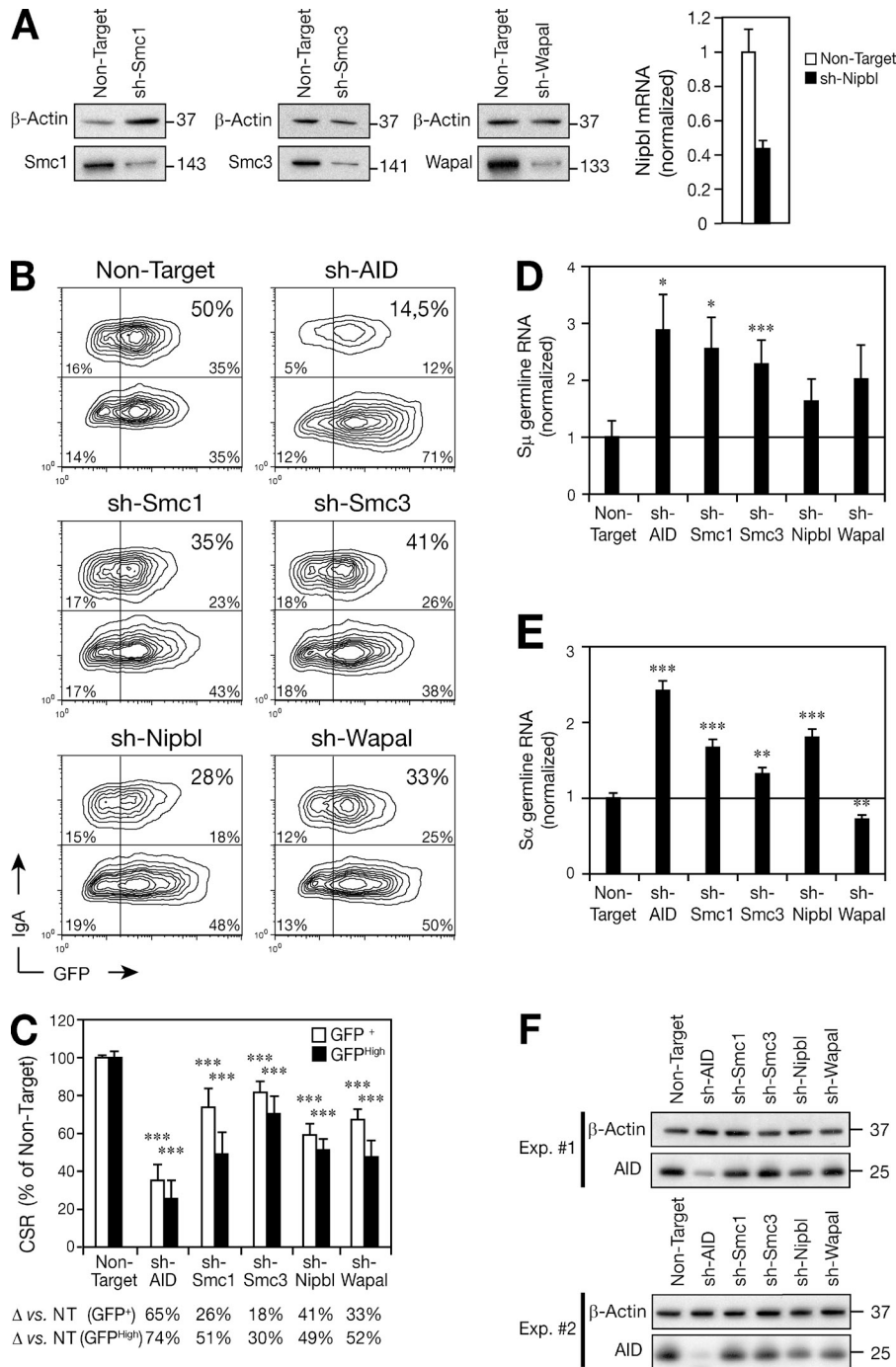


Figure 3. CSR is impaired by the knock-down of cohesin subunits. (A) CH12 cells were transduced with a lentivirus expressing a GFP reporter and shRNAs specific for AID, Smc1, Smc3, Nipbl, Wapal, or a Non-Target control. Transduced cells were stimulated for 48 h and sorted for GFP expression. Protein extracts and cDNAs were prepared and knock-down was determined by Western blotting or qPCR. Western blot for β -actin, Smc1, Smc3, and Wapal and qRT-PCR for *Nipbl* transcripts are shown. Expression was normalized to *Cd79b* and is presented relative to the Non-Target control, set as 1. Mean of triplicate samples (+SD) is shown. Statistical significance versus the Non-Target control (two-tailed Student's *t* test): $P = 0.0023$. Data are representative of three experiments. (B) CH12 cells treated as in A were analyzed for surface IgA and GFP expression by flow cytometry. Representative plots from four to eight independent experiments are shown. (C) CH12 cells treated as in A were gated on cells expressing GFP (GFP⁺; white bars) or high levels of GFP (GFP^{High}; black bars). The percentage (+SD) of CSR relative to the Non-Target shRNA control from four to eight independent experiments is shown. CSR in cells expressing the Non-Target shRNA control was set to 100%. The difference in CSR efficiency relative to the Non-Target control (Δ) is indicated below. Statistical significance versus the Non-Target control (two-tailed Student's *t* test) is indicated: ***, $P \leq 0.001$. (D and E) cDNA was prepared from CH12 cells treated as in A and qRT-PCR for μ (D) and α (E) germline transcripts was performed. Expression was normalized to *HPRT* mRNA abundance and is presented relative to the Non-Target control, set as 1 (black line). Mean of triplicate samples (+SD) is shown. Statistical significance versus the Non-Target control (two-tailed Student's *t* test) is indicated: *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$. (F) Proteins extracts were prepared from CH12 cells treated as in A. Western blots for β -actin and AID are shown. Data are representative of three independent experiments. Theoretical molecular masses in kilodaltons are indicated.

the IgH locus (Fig. 2, C and D). We conclude that Smc1 and Smc3 are dynamically recruited, independently of CTCF, to the IgH locus (at the S μ -C μ region) during CSR. As E μ is not bound by cohesin in resting B cells, the constitutive long-range interactions between E μ and the 3'RR that take place in resting B cells (Wuerffel et al., 2007) are most likely cohesin-independent. Nevertheless, given the dynamic recruitment of Smc1 and Smc3 at S μ -C μ (and possibly S γ 1) in activated B cells, we speculate that cohesin may play a role in supporting the structural changes occurring at the IgH locus upon B cell activation.

Cohesin is required for efficient CSR

To determine the functional relevance of the cohesin complex in CSR, we undertook knockdown experiments in CH12 cells, a B cell line which can be induced to undergo CSR from IgM to IgA in vitro and which allows the study of the role of specific factors in CSR (Pavri et al., 2010; Willmann et al., 2012). CH12 cells were transduced with lentiviruses expressing a GFP reporter together with shRNAs specific for AID (as a positive control), the core subunits of the cohesin complex (Smc1 and Smc3), the cohesin loader/unloader

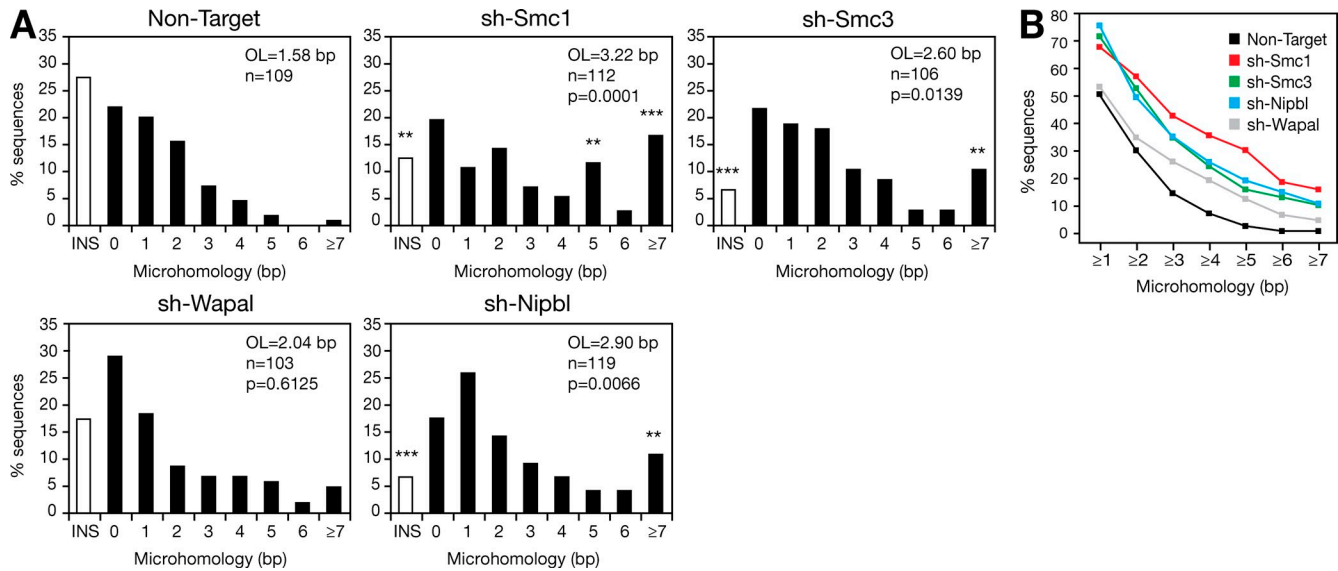


Figure 4. Knockdown of cohesin affects NHEJ. (A) CH12 cells were transduced with a lentivirus expressing a GFP reporter and shRNAs specific for AID, Smc1, Smc3, Nipbl, Wapal, or a Non-Target control. 48 h after stimulation, GFP-expressing cells were sorted. $S\mu$ - $S\alpha$ switch junctions were amplified by PCR, cloned, and sequenced. Bar graphs show the percentage of switch junction sequences with indicated nucleotide overlap. Number of junctions analyzed (n), mean length of overlap (OL), and p-values relative to the Non-Target control (Mann-Whitney test) are indicated. White bars indicate the percentage of sequences with small (1–4 nucleotides) insertions. Overlap was determined by identifying the longest region of perfect uninterrupted donor/acceptor identity. Sequences with insertions were not included in the calculation of the mean length of overlap. Significant differences relative to the Non-Target control (χ^2 test) are indicated: **, $P \leq 0.01$; ***, $P \leq 0.0001$. Data are from three independent experiments. (B) Cumulative percentage of sequences with a given length of microhomology (bp) and obtained from CH12 cells transduced with lentiviruses expressing shRNAs specific for Smc1 (red squares), Smc3 (green squares), Nipbl (blue squares), Wapal (gray squares), or a Non-Target negative control (black squares) and sorted for GFP expression. Data are from three independent experiments.

subunits (Nipbl and Wapal), and a Non-Target shRNA as a negative control. Knockdown efficiencies were determined by Western blotting or by quantitative RT-PCR (qRT-PCR) on GFP⁺ sorted cells (Fig. 3 A). Transduced cells were stimulated for 48 h, and their ability to undergo CSR to IgA was determined by flow cytometry (Fig. 3, B and C). As expected, knockdown of AID resulted in a robust reduction in the efficiency of CSR relative to the Non-Target shRNA control (Fig. 3, B and C). Interestingly, we found that knockdown of Smc1, Smc3, Nipbl, and Wapal resulted in a significant reduction in the efficiency of CSR (18–41%) in GFP⁺ cells (Fig. 3, B and C). This reduction was more pronounced (30–52%) when the analysis was performed by gating on cells expressing high levels of GFP (Fig. 3 C). The effect on CSR after cohesin knockdown was not due to decreased survival (Topro-3 staining; unpublished data), strong defects in proliferation (CFSE dilution; unpublished data), significant activation of the DNA damage response and cell cycle checkpoints (Western blot for γ -H2AX and p-Chk1; unpublished data), or defective cell cycle progression (flow cytometry; unpublished data).

To determine whether switch region transcription is affected by the knockdown of cohesin subunits, we measured the level of donor ($S\mu$) and acceptor ($S\alpha$) switch region transcripts by qRT-PCR in activated CH12 cells. We found that the level of $S\mu$ and $S\alpha$ transcripts was increased after knockdown of AID and cohesin (relative to the Non-Target control), with the exception of $S\alpha$ transcripts after knockdown of

Wapal (Fig. 3, D and E), as expected from cells in which CSR is compromised and that continue to transcribe the switch regions. As no significant reduction in the level of these transcripts after Smc1, Smc3, and Nipbl knockdown was observed, we conclude that switch regions continue to be efficiently transcribed and that they are accessible for DNA deamination by AID. Therefore, cohesin appears not to be involved in the transcriptional regulation of switch regions during CSR. Importantly, we excluded a potential reduction in AID expression levels by Western blot (Fig. 3 F). We conclude that the cohesin complex is required for efficient CSR in CH12 cells. The role of cohesin in CSR appears to be independent of regulating switch region transcription and/or AID accessibility. Concerning a potential more global effect on transcription, we cannot exclude the possibility that the expression of additional genes required for CSR (other than AID) is affected by the knockdown of cohesin.

Knockdown of cohesin affects NHEJ

DSBs triggered by AID in switch regions during CSR are resolved through the NHEJ pathway, and the resulting switch junctions display small insertions and short stretches of microhomology (Stavnezer et al., 2010). In the absence of core NHEJ components, an increase in the usage of microhomology is observed concomitantly with a complete loss of direct joining (Yan et al., 2007). To determine whether cohesin knockdown affects the resolution of DSBs generated during

CSR, we cloned and sequenced $\Sigma\mu/\Sigma\alpha$ switch junctions from stimulated CH12 transduced with lentiviruses expressing shRNAs for Smc1, Smc3, Nipbl, Wapal, and a Non-Target negative control (Fig. 4) and sorted for GFP expression. Sequence analysis (Stavnezer et al., 2010) revealed that knockdown of cohesin subunits resulted in a significant increase in the usage of microhomology when compared with the Non-Target control (Fig. 4). Although the mean length of overlap (excluding insertions) was of 1.58 bp for the Non-Target control, it was increased to 3.22 bp for Smc1 ($P = 0.0001$), 2.60 bp for Smc3 ($P = 0.0139$), and 2.90 bp for Nipbl ($P = 0.0066$). The switch junctions obtained after Wapal knockdown displayed an overlap of 2.04 bp that was not statistically different from the Non-Target control ($P = 0.6125$). The increase in microhomology was due to sequences bearing >7 bp of microhomology at the junction and a reduction in those bearing short insertions (Fig. 4), similar to what has been described in human patients with deficiency in DNA ligase IV (Du et al., 2008), Artemis (Du et al., 2008), or ATM (Pan-Hammarström et al., 2006). In contrast to deficiency in core NHEJ components (Stavnezer et al., 2010), we did not find a reduction in the frequency of direct joining events (Fig. 4). We conclude that switch recombination junctions generated after Smc1, Smc3, and Nipbl knockdown (but not Wapal) are biased toward the usage of longer microhomologies. Given the role of Wapal in releasing cohesin from chromatin (Kueng et al., 2006), this suggests that cohesin is recruited but not released from the IgH locus and that NHEJ proceeds unaffected. Therefore, it appears that the loading of cohesin is sufficient to determine the outcome of DSB repair and that cohesin participates in the resolution of AID-induced DNA breaks.

Increased usage of microhomology at the junctions is reminiscent of what is observed in B cells defective for core components of the NHEJ pathway (Yan et al., 2007). Nevertheless, deficiency in XRCC4 or DNA ligase IV also results in a complete loss of sequences repaired through a direct joining (Yan et al., 2007). Therefore, it is unlikely that the cohesin complex is, by itself, part of the NHEJ machinery. As cohesin has been implicated in the recruitment of 53BP1 to γ -irradiation-induced foci (Watrinn and Peters, 2009) and 53BP1 deficiency leads to defective CSR, increased DNA end resection, and preferential usage of microhomology (Bothmer et al., 2010), we speculate that cohesin could participate in the recruitment of 53BP1 to AID-induced DSBs and that defective 53BP1 recruitment could account for the increased usage of microhomology observed. Overall, our results implicate the cohesin complex in the mechanism of CSR and provide evidence for its involvement in regulating the repair of programmed DSBs.

MATERIALS AND METHODS

Nuclear extracts and coimmunoprecipitation. Nuclear extracts and chromatin fractions were prepared using standard techniques (in the presence of 100 U/ml Benzonase; Novagen) from CH12F3 cells stably expressing AID^{FLAG-HA}, EGFP^{FLAG-HA}, or the tags alone (Jeevan-Raj et al., 2011). Coimmunoprecipitations (in the presence of 100 U/ml Benzonase; Novagen) and Western blot analysis were performed as previously described (Jeevan-Raj et al., 2011). See Table S3 for antibodies used.

Mass spectrometry analysis. 20 mg nuclear extract were immunoprecipitated with Flag M2-agarose beads, washed, and eluted with Flag peptide as previously described (Jeevan-Raj et al., 2011). Flag eluates were fractionated by one-dimensional electrophoresis and processed as previously described (Jeevan-Raj et al., 2011) for identification by nano-LC-MS/MS or directly submitted to Multidimensional Protein Identification Technology (MudPIT). MudPIT analyses were performed as previously described (Washburn et al., 2001; Florens et al., 2006). In brief, protein mixtures were TCA-precipitated, urea-denatured, reduced, alkylated, and digested with endoproteinase Lys-C (Roche), followed by modified trypsin digestion (Promega). Peptide mixtures were loaded onto a triphasic 100 μ m inner diameter fused silica microcapillary column. Loaded columns were placed in-line with a Dionex Ultimate 3000 nano-LC (Thermo Fisher Scientific) and an LTQ Velos linear ion trap mass spectrometer equipped with a nano-LC electrospray ionization source (Thermo Fisher Scientific). A fully automated 12-step MudPIT run was performed as previously described (Florens et al., 2006), during which each full MS scan (from 300 to 1,700 m/z range) was followed by 20 MS/MS events using data-dependent acquisition. Proteins were identified by database searching using SEQUEST (Thermo Fisher Scientific) with Proteome Discoverer 1.3 software (Thermo Fisher Scientific) against the mouse Swissprot database (2011-02 release). Peptides were filtered with Xcorr versus charge state 1.5–1, 2.5–2, 3–3, 3.2–4, and peptides of at least 7 amino acids in length.

shRNA-mediated knockdown. The lentiviral vectors (pLKO.1 and pLKO.1-puro-CMV-TurboGFP) expressing shRNAs specific for AID (TRCN0000112031), Smc1 (TRCN0000109034), Smc3 (TRCN0000109007), Nipbl (TRCN0000124037), and Wapal (TRCN0000177268) or a Non-Target control (SHC002) were obtained from Sigma-Aldrich. The lentiviral vectors were transiently transfected into Lenti-X 293T cells (Takara Bio Inc.) to produce infectious viral particles as previously described (Willmann et al., 2012). 2 d later, CH12 cells were spin-infected with viral supernatants supplemented with 10 μ g/ml polybrene (Sigma-Aldrich). Cells were selected for 5 d with 1 μ g/ml puromycin before CSR induction.

Real-time quantitative (q) RT-PCR. RNA and cDNA were prepared using standard techniques. qPCR was performed in triplicates using the Universal Probe Library (UPL) system (Roche) or SyberGreen (QIAGEN) and a LightCycler 480 (Roche). Transcript quantities were calculated relative to standard curves and normalized to β -actin, CD79b, or HPRT mRNA. See Table S4 for primers and probes.

Cell culture and flow cytometry. Lentivirally transduced CH12 cells were cultured with 5 ng/ml IL-4 (Sigma-Aldrich), 1 ng/ml TGF- β (R&D System), 200 ng/ml monoclonal anti-CD40 antibody (eBioscience), and 1 μ g/ml puromycin and analyzed after 48–72 h for cell surface expression of IgA by flow cytometry as previously described (Robert et al., 2009). Resting splenic B cells were isolated from 8–12-wk-old C57BL/6 mice using CD43 microbeads (Miltenyi Biotec) and cultured for 60 h with 50 μ g/ml LPS (Sigma-Aldrich) and 5 ng/ml IL-4 (Sigma-Aldrich) as previously described (Jeevan-Raj et al., 2011). All animal work was performed under protocols approved by the Direction des Services Vétérinaires du Bas-Rhin, France (Authorization No. 67–343).

Switch junction analysis. $\Sigma\mu$ - $\Sigma\alpha$ switch junctions were amplified using previously described primers (Ehrenstein and Neuberger, 1999; Schrader et al., 2002) and conditions (Robert et al., 2009) from genomic DNA prepared from lentivirally transduced CH12 cells stimulated for 72 h and sorted for GFP expression. PCR products were cloned using TOPO-TA cloning kit (Invitrogen) and sequenced using T7 universal primer. Sequence analysis was performed as previously described (Robert et al., 2009).

ChIP-Seq. Resting or activated B cells were cross-linked for 10 min at 37°C with 1% (vol/vol) formaldehyde, followed by quenching with glycine (0.125 M final concentration). Cross-linked samples were then sonicated to obtain DNA fragments 200–500 bp in length using a sonicator (Covaris).

Chromatin (from 10×10^7 cells) was precleared with protein A magnetic beads prewashed with PBS 0.05% tween, 5% BSA, and immunoprecipitated in ChIP dilution buffer (0.06% SDS, 20 mM Tris, pH 8.1, 2 mM EDTA, 160 mM NaCl, 1.045% Triton X-100, and 0.05% proteinase inhibitor cocktail) overnight at 4°C with protein A magnetic beads (Invitrogen) coupled to 100 µg Smc1 or Smc3 antibodies and processed according to the Millipore protocol. Cross-links were reversed for 4 h at 65°C in Tris-EDTA buffer with 0.3% (wt/vol) SDS and 1 mg/ml proteinase K. ChIP DNA was extracted with an IPure kit (Diagenode). Libraries were prepared for sequencing according to the manufacturer's protocol (Illumina) and sequenced on the Illumina Genome Analyzer IIx as single-end 50 base reads according to manufacturer instructions. Image analysis and base calling were performed using the Illumina Pipeline and sequence reads were mapped to reference genome mm9/NCBI37 using Bowtie v0.12.7. Peak calling was performed using MACS (Zhang et al., 2008) with default parameters. Global comparison of samples and clustering analysis were performed using seqMINER (Ye et al., 2011).

ChIP-qPCR. Analytical-scale ChIP was performed on chromatin prepared from 10^7 (resting or activated) splenic B cells isolated from a pool of five mice. qPCR was performed at several locations across the IgH locus using primers listed in Table S4. Results are expressed as percent input and represent the mean of three qPCR technical replicates. Error bars represent the corresponding standard deviation.

Accession codes. ChIP-Seq data for CTCF, Smc1, and Smc3 on resting and activated B cells was submitted to GEO (GSE43594).

Online supplemental material. Table S1 lists known AID partner proteins found. Table S2 lists novel AID partner proteins found. Table S3 lists antibodies. Table S4 lists primers. Table S5 shows ChIP-qPCR statistical analysis. Online supplemental material is available at <http://www.jem.org/cgi/content/full/jem.20130166/DC1>.

We thank members of the Reina-San-Martin laboratory for discussions; E. Soutoglou, I. Sumara, and M. Nussenzweig for comments on the manuscript; S. Pattabhiraman, M. Mendoza Parra, S. Legras, and B. Jost for assistance with ChIP-Seq analyses; V. Chavant and F. Ruffenach for assistance in mass spectrometry analysis; C. Ebel for cell sorting; and A. Gazumyan for advice on lentiviral infections.

A.-S. Thomas-Claudepierre was supported by the Ministère de l'Enseignement Supérieur et de la Recherche, France and the Fondation ARC. E. Schiavo was supported by the IGBMC International PhD Program and the Fondation ARC. This work was supported by grants to B. Reina-San-Martin from the Agence Nationale pour la Recherche (ANR-Blanc), the Fondation ARC (Program ARC), and the Institut National de la Santé et de la Recherche Médicale (Avenir-INSERM).

The authors have no conflicting financial interests.

Submitted: 23 January 2013

Accepted: 29 August 2013

REFERENCES

- Basu, U., F.L. Meng, C. Keim, V. Grinstein, E. Pefanis, J. Eccleston, T. Zhang, D. Myers, C.R. Wasserman, D.R. Wesemann, et al. 2011. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell*. 144:353–363. <http://dx.doi.org/10.1016/j.cell.2011.01.001>
- Bothmer, A., D.F. Robbiani, N. Feldhahn, A. Gazumyan, A. Nussenzweig, and M.C. Nussenzweig. 2010. 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. *J. Exp. Med.* 207:855–865. <http://dx.doi.org/10.1084/jem.20100244>
- Chatterjee, S., Z. Ju, R. Hassan, S.A. Volpi, A.V. Emelyanov, and B.K. Birshtein. 2011. Dynamic changes in binding of immunoglobulin heavy chain 3' regulatory region to protein factors during class switching. *J. Biol. Chem.* 286:29303–29312. <http://dx.doi.org/10.1074/jbc.M111.243543>
- Chaudhuri, J., U. Basu, A. Zarrin, C. Yan, S. Franco, T. Perlot, B. Vuong, J. Wang, R.T. Phan, A. Datta, et al. 2007. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv. Immunol.* 94:157–214. [http://dx.doi.org/10.1016/S0065-2776\(06\)94006-1](http://dx.doi.org/10.1016/S0065-2776(06)94006-1)
- Degner, S.C., J. Verma-Gaur, T.P. Wong, C. Bossen, G.M. Iverson, A. Torkamani, C. Vettermann, Y.C. Lin, Z. Ju, D. Schulz, et al. 2011. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the IgH locus and antisense transcription in pro-B cells. *Proc. Natl. Acad. Sci. USA*. 108:9566–9571. <http://dx.doi.org/10.1073/pnas.1019391108>
- Di Noia, J.M., and M.S. Neuberger. 2007. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* 76:1–22. <http://dx.doi.org/10.1146/annurev.biochem.76.061705.090740>
- Dorsett, D. 2009. Cohesin, gene expression and development: lessons from *Drosophila*. *Chromosome Res.* 17:185–200. <http://dx.doi.org/10.1007/s10577-009-9022-5>
- Du, L., M. van der Burg, S.W. Popov, A. Kotnis, J.J. van Dongen, A.R. Gennery, and Q. Pan-Hammarström. 2008. Involvement of Artemis in nonhomologous end-joining during immunoglobulin class switch recombination. *J. Exp. Med.* 205:3031–3040. <http://dx.doi.org/10.1084/jem.20081915>
- Ehrenstein, M.R., and M.S. Neuberger. 1999. Deficiency in Msh2 affects the efficiency and local sequence specificity of immunoglobulin class-switch recombination: parallels with somatic hypermutation. *EMBO J.* 18:3484–3490. <http://dx.doi.org/10.1093/emboj/18.12.3484>
- Florens, L., M.J. Carozza, S.K. Swanson, M. Fournier, M.K. Coleman, J.L. Workman, and M.P. Washburn. 2006. Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods*. 40:303–311. <http://dx.doi.org/10.1016/j.jymeth.2006.07.028>
- Guo, C., H.S. Yoon, A. Franklin, S. Jain, A. Ebert, H.L. Cheng, E. Hansen, O. Despo, C. Bossen, C. Vettermann, et al. 2011. CTCF-binding elements mediate control of V(D)J recombination. *Nature*. 477:424–430. <http://dx.doi.org/10.1038/nature10495>
- Jeevan-Raj, B.P., I. Robert, V. Heyer, A. Page, J.H. Wang, F. Cammas, F.W. Alt, R. Losson, and B. Reina-San-Martin. 2011. Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *J. Exp. Med.* 208:1649–1660. <http://dx.doi.org/10.1084/jem.20110118>
- Kagey, M.H., J.J. Newman, S. Bilodeau, Y. Zhan, D.A. Orlando, N.L. van Berkum, C.C. Ebmeier, J. Goossens, P.B. Rahl, S.S. Levine, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature*. 467:430–435. <http://dx.doi.org/10.1038/nature09380>
- Kenter, A.L., S. Feldman, R. Wuerrfel, I. Achour, L. Wang, and S. Kumar. 2012. Three-dimensional architecture of the IgH locus facilitates class switch recombination. *Ann. N. Y. Acad. Sci.* 1267:86–94. <http://dx.doi.org/10.1111/j.1749-6632.2012.06604.x>
- Kueng, S., B. Hegemann, B.H. Peters, J.J. Lipp, A. Schleiffer, K. Mechtler, and J.M. Peters. 2006. Wapl controls the dynamic association of cohesin with chromatin. *Cell*. 127:955–967. <http://dx.doi.org/10.1016/j.cell.2006.09.040>
- Nakahashi, H., K.R. Kwon, W. Resch, L. Vian, M. Dose, D. Stavreva, O. Hakim, N. Pruett, S. Nelson, A. Yamane, et al. 2013. A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Rep.* 3:1678–1689. <http://dx.doi.org/10.1016/j.celrep.2013.04.024>
- Nasmyth, K., and C.H. Haering. 2009. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* 43:525–558. <http://dx.doi.org/10.1146/annurev-genet-102108-134233>
- Pan-Hammarström, Q., A. Lähdesmäki, Y. Zhao, L. Du, Z. Zhao, S. Wen, V.L. Ruiz-Perez, D.K. Dunn-Walters, J.A. Goodship, and L. Hammarström. 2006. Disparate roles of ATR and ATM in immunoglobulin class switch recombination and somatic hypermutation. *J. Exp. Med.* 203:99–110. <http://dx.doi.org/10.1084/jem.20050595>
- Pavri, R., A. Gazumyan, M. Jankovic, M. Di Virgilio, I. Klein, C. Ansarah-Sobrinho, W. Resch, A. Yamane, B. Reina-San-Martin, V. Barreto, et al. 2010. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell*. 143:122–133. <http://dx.doi.org/10.1016/j.cell.2010.09.017>
- Petersen-Mahrt, S.K., R.S. Harris, and M.S. Neuberger. 2002. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature*. 418:99–103. <http://dx.doi.org/10.1038/nature00862>

- Ramiro, A., B. Reina San-Martin, K. McBride, M. Jankovic, V. Barreto, A. Nussenzweig, and M.C. Nussenzweig. 2007. The role of activation-induced deaminase in antibody diversification and chromosome translocations. *Adv Immunol.* 94:75–107. [http://dx.doi.org/10.1016/S0065-2776\(06\)94003-6](http://dx.doi.org/10.1016/S0065-2776(06)94003-6)
- Robert, I., F. Dantzer, and B. Reina-San-Martin. 2009. Parp1 facilitates alternative NHEJ, whereas Parp2 suppresses IgH/c-myc translocations during immunoglobulin class switch recombination. *J. Exp. Med.* 206:1047–1056. <http://dx.doi.org/10.1084/jem.20082468>
- Schrader, C.E., J. Vardo, and J. Stavnezer. 2002. Role for mismatch repair proteins Msh2, Mlh1, and Pms2 in immunoglobulin class switching shown by sequence analysis of recombination junctions. *J. Exp. Med.* 195:367–373. <http://dx.doi.org/10.1084/jem.20011877>
- Seitan, V.C., B. Hao, K. Tachibana-Konwalski, T. Lavagnoli, H. Mira-Bontenbal, K.E. Brown, G. Teng, T. Carroll, A. Terry, K. Horan, et al. 2011. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature.* 476:467–471. <http://dx.doi.org/10.1038/nature10312>
- Stavnezer, J., A. Björkman, L. Du, A. Cagigi, and Q. Pan-Hammarström. 2010. Mapping of switch recombination junctions, a tool for studying DNA repair pathways during immunoglobulin class switching. *Adv Immunol.* 108:45–109. <http://dx.doi.org/10.1016/B978-0-12-380995-7.00003-3>
- Washburn, M.P., D. Wolters, and J.R. Yates III. 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19:242–247. <http://dx.doi.org/10.1038/85686>
- Watrin, E., and J.M. Peters. 2009. The cohesin complex is required for the DNA damage-induced G2/M checkpoint in mammalian cells. *EMBO J.* 28:2625–2635. <http://dx.doi.org/10.1038/emboj.2009.202>
- Willmann, K.L., S. Milosevic, S. Pauklin, K.M. Schmitz, G. Rangam, M.T. Simon, S. Maslen, M. Skehel, I. Robert, V. Heyer, et al. 2012. A role for the RNA pol II-associated PAF complex in AID-induced immune diversification. *J. Exp. Med.* 209:2099–2111. <http://dx.doi.org/10.1084/jem.20112145>
- Wuerffel, R., L. Wang, F. Grigera, J. Manis, E. Selsing, T. Perlot, F.W. Alt, M. Cogne, E. Pinaud, and A.L. Kenter. 2007. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity.* 27:711–722. <http://dx.doi.org/10.1016/j.immuni.2007.09.007>
- Yamane, A., W. Resch, N. Kuo, S. Kuchen, Z. Li, H.W. Sun, D.F. Robbiani, K. McBride, M.C. Nussenzweig, and R. Casellas. 2011. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat. Immunol.* 12:62–69. <http://dx.doi.org/10.1038/ni.1964>
- Yan, C.T., C. Boboila, E.K. Souza, S. Franco, T.R. Hickernell, M. Murphy, S. Gumaste, M. Geyer, A.A. Zarrin, J.P. Manis, et al. 2007. IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature.* 449:478–482. <http://dx.doi.org/10.1038/nature06020>
- Ye, T., A.R. Krebs, M.A. Choukallah, C. Keime, F. Plewniak, I. Davidson, and L. Tora. 2011. seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res.* 39:e35. <http://dx.doi.org/10.1093/nar/gkq1287>
- Zhang, Y., T. Liu, C.A. Meyer, J. Eeckhoutte, D.S. Johnson, B.E. Bernstein, C. Nusbaum, R.M. Myers, M. Brown, W. Li, and X.S. Liu. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137. <http://dx.doi.org/10.1186/gb-2008-9-9-r137>

Mediator facilitates transcriptional activation and dynamic long-range contacts at the IgH locus during class switch recombination

Anne-Sophie Thomas-Claudepierre,^{1,2,3,4} Isabelle Robert,^{1,2,3,4} Pedro P. Rocha,⁵ Ramya Raviram,^{5,6} Ebe Schiavo,^{1,2,3,4} Vincent Heyer,^{1,2,3,4} Richard Bonneau,^{6,8,9} Vincent M. Luo,^{5,6} Janardan K. Reddy,¹⁰ Tilman Borggrefe,¹¹ Jane A. Skok,^{5,7} and Bernardo Reina-San-Martin^{1,2,3,4}

¹Institut de Génétique et de Biologie Moléculaire et Cellulaire, 67400 Illkirch, France

²Institut National de la Santé et de la Recherche Médicale, Unité 964, 67404 Illkirch, France

³Centre National de la Recherche Scientifique, Unité Mixte de Recherche 7104, 67404 Illkirch, France

⁴Université de Strasbourg, 67400 Illkirch, France

⁵Department of Pathology, School of Medicine, ⁶Department of Biology, and ⁷New York University Cancer Institute, New York University, New York, NY 10003

⁸Department of Computer Science, Courant Institute of Mathematical Sciences, New York, NY 10003

⁹Simons Center for Data Analysis, New York, NY 10010

¹⁰Department of Pathology, Feinberg School of Medicine, Northwestern University, Chicago, IL 60208

¹¹Justus Liebig Universität Giessen, 35390 Giessen, Germany

Immunoglobulin (Ig) class switch recombination (CSR) is initiated by the transcription-coupled recruitment of activation-induced cytidine deaminase (AID) to Ig switch regions (S regions). During CSR, the IgH locus undergoes dynamic three-dimensional structural changes in which promoters, enhancers, and S regions are brought to close proximity. Nevertheless, little is known about the underlying mechanisms. In this study, we show that Med1 and Med12, two subunits of the mediator complex implicated in transcription initiation and long-range enhancer/promoter loop formation, are dynamically recruited to the IgH locus enhancers and the acceptor regions during CSR and that their knockdown in CH12 cells results in impaired CSR. Furthermore, we show that conditional inactivation of *Med1* in B cells results in defective CSR and reduced acceptor S region transcription. Finally, we show that in B cells undergoing CSR, the dynamic long-range contacts between the IgH enhancers and the acceptor regions correlate with Med1 and Med12 binding and that they happen at a reduced frequency in *Med1*-deficient B cells. Our results implicate the mediator complex in the mechanism of CSR and are consistent with a model in which mediator facilitates the long-range contacts between S regions and the IgH locus enhancers during CSR and their transcriptional activation.

Ig class switch recombination (CSR) is a long-range DNA recombination reaction that occurs between Ig switch regions (S regions) and that replaces the isotype expressed (from IgM to IgG, IgE, or IgA), providing novel effector functions for efficient antigen clearance (Chaudhuri et al., 2007). CSR is initiated by the transcription-coupled recruitment of activation-induced cytidine deaminase (AID; Basu et al., 2011; Pavri and Nussenzweig, 2011), an enzyme that deaminates cytosines into uracils in the single-strand DNA exposed by transcription (Petersen-Mahrt et al., 2002). During CSR, the choice of recombination to a particular isotype is determined by the activation of specific S region promoters (Basu et al., 2011; Pavri and Nussenzweig, 2011), triggering the generation of noncoding germline transcripts (Chaudhuri et al., 2007). Germline transcription precedes recombination, is induced

at both the donor and acceptor S regions, and is required for recombination (Chaudhuri et al., 2007). Transcriptional activation of the IgH locus during CSR is controlled by the E μ enhancer located upstream of the donor (S μ) S region and by a major regulatory region (RR) located at the 3' end of the locus (3' RR). Both of these enhancer elements are required for transcription and for CSR (Chaudhuri et al., 2007; Pavri and Nussenzweig, 2011). The current model is that during CSR, recombination between S regions proceeds by the inducible formation of long-range DNA loops involving the S region promoters and the E μ and 3' RR enhancers (Wuerffel et al., 2007; Kenter et al., 2012), possibly through transcription factors (Feldman et al., 2015). Nevertheless, the molecular mechanisms controlling these conformational changes remain to be elucidated.

Correspondence to Bernardo Reina-San-Martin: reinab@igbmc.fr

Abbreviations used: AID, activation-induced cytidine deaminase; ChIP, chromatin immunoprecipitation; CSR, class switch recombination; qPCR, quantitative PCR; RR, regulatory region; RT, real time; S region, switch region.

© 2016 Thomas-Claudepierre et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

Mediator is an evolutionarily conserved multiprotein complex composed of 31 subunits organized in four modules that is required for gene transcription by RNA polymerase II (Pol II; Malik and Roeder, 2010; Conaway and Conaway, 2011). The head, middle, and tail modules form a stable core complex that associates reversibly with the CDK8 module (consisting of cyclin-dependent kinase 8, cyclin C, Med12, and Med13) to control interactions of mediator with the Pol II machinery (Malik and Roeder, 2010; Conaway and Conaway, 2011). Mediator behaves as an interface between Pol II and transcription factors and is capable of promoting Pol II preinitiation complex assembly, transcription initiation by Pol II, regulation of Pol II pausing and elongation, recruitment of transcription elongation factors, and control of the phosphorylation state of the C-terminal domain of Pol II (Malik and Roeder, 2010; Conaway and Conaway, 2011; Allen and Taatjes, 2015). The Med1 subunit of mediator, part of the middle module, interacts with distinct transcriptional activators (Borggreffe and Yue, 2011) and has been shown to play a key role in embryonic development (Ito et al., 2000; Zhu et al., 2000), erythropoiesis (Stumpf et al., 2010), and iNKT cell development (Yue et al., 2011). In addition, Med1 recruitment to chromatin is one of the features that characterizes super enhancers (Whyte et al., 2013). Interestingly, mediator has also been implicated, together with cohesin, in the formation of long-range DNA loops (Malik and Roeder, 2010; Conaway and Conaway, 2011; Allen and Taatjes, 2015), and chromatin immunoprecipitation sequencing (ChIP-Seq) analysis for Smc1, Smc3, Med1, and Med12 revealed that cohesin–mediator binding predicts genomic sites of long-range promoter–enhancer interactions (Kagey et al., 2010; Phillips-Cremins et al., 2013). As we have recently implicated the cohesin complex in the mechanism of CSR (Thomas-Claudepierre et al., 2013), we have examined the role of mediator in CSR by performing shRNA-mediated knockdowns of the Med1 and Med12 subunits of mediator (belonging to different modules) in CH12 cells and by conditionally inactivating the Med1 subunit in developing B cells.

RESULTS AND DISCUSSION

The mediator complex is dynamically recruited to the IgH locus

Mediator has been suggested to generate DNA loops by binding to promoters and enhancers in embryonic stem cells to induce gene expression (Kagey et al., 2010; Phillips-Cremins et al., 2013). Recently, a ChIP-Seq experiment for the Med12 subunit of mediator performed on ex vivo–activated WT B cells has been reported (Wang et al., 2014). Analysis of the available data (SRR975483) showed that in B cells induced to switch to IgG1 (LPS + IL-4), Med12 is significantly recruited at E μ , the 3' RR (hs1,2 and hs4), the γ 1 promoter (γ 1p), the S γ 1 S region, and a region downstream of C γ 1 containing a putative enhancer (γ 1E; Medvedovic et al., 2013) that is also bound by Med1 in developing B cells (Whyte et al., 2013; Predeus et al., 2014).

To confirm these results and investigate the dynamics of mediator binding during CSR, we performed ChIP–quantitative PCR (qPCR) on resting B cells and B cells stimulated with LPS + IL-4 or LPS from control (*Med1*^{F/F}) and *Med1*-deficient mice (*Med1*^{F/F}*Mb1*^{Cre/+}; see section CSR and transcription S regions are compromised by Med1 deficiency in primary B cells) using Med1- and Med12-specific antibodies. We used primer pairs spanning across the IgH locus: E μ , the S region promoters (γ 3p, γ 1p, γ 2bp), the γ 1E enhancer, two elements of the 3' RR bound by Med12 (hs1,2 and hs4), and C γ 2a as negative control. In control resting B cells, we found that Med1 and Med12 are exclusively recruited at the E μ enhancer and at the hs1,2 and hs4 elements of the 3' RR (Fig. 1 A). Binding was specific, as the ChIP signal was significantly enriched relative to *Med1*-deficient B cells. Upon CSR induction to IgG1 with LPS + IL-4, Med1 and Med12 are in addition recruited to γ 1p and γ 1E but not to γ 3p and γ 2p (Fig. 1 B and Table S1). Conversely, stimulation with LPS, which induces CSR to IgG3 and IgG2b, leads to the specific recruitment of Med1 and Med12 to γ 3p, γ 2bp, and γ 1E, but not to γ 1p (Fig. 1 C and Table S1). Additionally, ChIP–qPCR for Med12 performed on *Med1*-deficient cells did not show reduced binding (Fig. 1, A–C; and Table S1), indicating that recruitment of Med12 is not dependent on Med1.

The Med1 and Med12 subunits, which belong to different modules of mediator, show similar recruitment profiles, suggesting that they are part of the same complex in B cells undergoing CSR. Consistent with this, we could show that Med1 and Med12 reciprocally coimmunoprecipitate from extracts prepared from both CH12 cells and ex vivo–activated splenic B cells (Fig. 1, D and E). We conclude that the mediator complex is dynamically recruited to the IgH locus at the promoters driving the transcription of acceptor S regions in a stimulation-specific manner and that Med12 can be recruited in the absence of Med1.

Knockdown of Med1 or Med12 impairs CSR and transcription of the acceptor S region in CH12 cells

To investigate the functional significance of Med1 and Med12 recruitment to the IgH locus during CSR, we undertook knockdown experiments in CH12 cells, which can be efficiently stimulated to undergo CSR from IgM to IgA in vitro (Thomas-Claudepierre et al., 2013). CH12 cells were transduced with lentiviruses expressing shRNAs specific for Med1 or Med12. As controls, we used a nontarget shRNA and AID-specific shRNA (Thomas-Claudepierre et al., 2013). Knockdown efficiency was confirmed by Western blot (Fig. 2 A), and the ability of transduced cells to undergo CSR to IgA was determined by flow cytometry (Fig. 2, B and C). As expected, knockdown of AID resulted in a robust reduction in the efficiency of CSR (65%) relative to the nontarget shRNA control (Fig. 2, B and C). Similarly, knockdown of the Med1 or Med12 subunits resulted in a significant reduction in the efficiency of CSR (37%) relative to the nontarget shRNA (Med1 vs. nontarget: $P = 3.66 \times 10^{-15}$; Med12

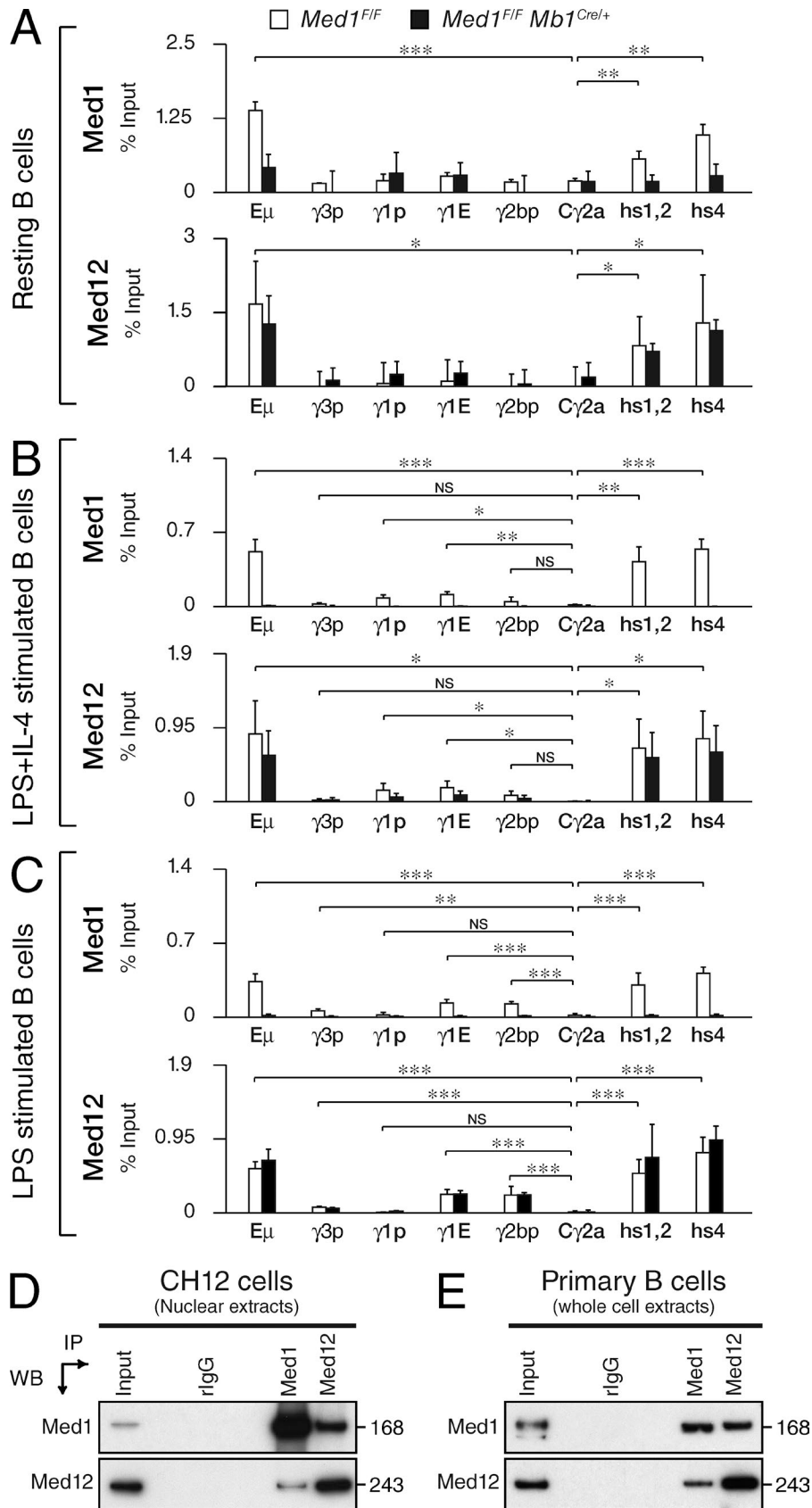


Figure 1. Med1 and Med12 are dynamically recruited to the IgH locus during CSR. (A–C) Med1 and Med12 ChIP–qPCR experiments performed on chromatin prepared from 10^7 splenic resting B cells (A) or cells stimulated with LPS + IL-4 (B) or LPS alone (C) for 72 h. qPCR was performed at the indicated locations (see Table S3 for primers). Results are expressed as the mean of the percentage input (+SD) from three independent biological replicate experiments. Statistical significance from WT results versus Cγ2a (two-tailed Student's *t* test) is indicated. *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$. White bars represent ChIP performed on WT samples; black bars represent ChIP performed on *Med1*-deficient samples. See Table S1 for full statistical analysis. (D and E) Nuclear extracts (D) from CH12 cells and whole-cell extracts (E) prepared from WT splenic B cells stimulated with LPS + IL-4 for 72 h were immunoprecipitated and blotted with antibodies specific for Med1 and Med12. Input represents 1% of material used. Theoretical molecular masses in kilodaltons are indicated. Data are representative of two independent experiments.

vs. nontarget: $P = 4.03 \times 10^{-10}$; Fig. 2, B and C). The CSR defect observed was independent of abnormalities in AID expression as determined by Western blot (Fig. 2 A). We conclude that knockdown of Med1 and Med12 in CH12 cells results in defective CSR without affecting AID expression. To determine whether Med1 and Med12 knockdown affects S region transcription, we analyzed the level of donor (I μ -C μ) and acceptor (I α -C α) germline transcripts by real time (RT)-qPCR in CH12 cells expressing shRNAs specific for Med1, Med12, or a nontarget control (Fig. 2, D and E). Although the level of donor germline transcript was not affected (Fig. 2 D), we found that the level of the acceptor S region germline transcripts was significantly reduced upon Med1 and Med12 knockdown when compared with the nontarget control (Fig. 2 E). We conclude that Med1 and Med12 are required for robust transcription of the S α acceptor S region.

CSR and transcription of acceptor S regions are compromised by *Med1* deficiency in primary B cells

To inactivate the Med1 subunit in developing B cells, we bred *Med1*-floxed mice (*Med1*^{F/F}; Jia et al., 2004) with *Mb1*^{Cre/+} knock-in mice (Hobeika et al., 2006). Despite efficient Cre-mediated deletion (not depicted), normal B cell numbers (not depicted) and frequencies were found in the bone marrow and the spleen (not depicted). The only difference observed was an increase in the proportion of marginal zone relative to follicular B cells in the spleen of *Med1*-deficient mice (not depicted). Therefore, mature B cell populations in the spleen are efficiently generated in the absence of Med1.

To determine whether conditional inactivation of *Med1* results in defective CSR, we cultured in vitro CFSE-labeled splenic B cells isolated from *Med1*^{F/F}*Mb1*^{Cre/+} mice and control mice (*Med1*^{+/+}, *Mb1*^{Cre/+}, *Med1*^{F/F}, or *Med1*^{F/+}*Mb1*^{Cre/+}) under conditions that induce CSR to different isotypes (Fig. 3, A and B). Although no differences were observed among the different control genotypes, interestingly, we found that *Med1* deficiency resulted in a 30–60% reduction in CSR to all isotypes tested (Fig. 3, A and B). To determine whether *Med1* deficiency affects AID expression, we measured the level of AID mRNA and protein in activated *Med1*^{F/F}*Mb1*^{Cre/+} and control B cells by RT-qPCR and Western blot (Fig. 3 C). We did not find any significant reduction in AID expression level in *Med1*^{F/F}*Mb1*^{Cre/+} mice compared with control mice (Fig. 3 C). Therefore, reduced CSR in *Med1*-deficient B cells cannot be explained by defective AID expression. In addition, the effect of *Med1* deficiency on CSR was not caused by decreased survival (not depicted), strong proliferation defects (not depicted), or defective cell cycle progression (not depicted), nor by an increased proportion of marginal zone B cells in *Med1*^{F/F}*Mb1*^{Cre/+} mice (not depicted). We conclude that *Med1* deletion results in a B cell-intrinsic CSR defect that is independent of defective AID expression or strong proliferation abnormalities.

To determine whether *Med1* deletion has a general impact on S region transcription, we measured the level of donor

(I μ -C μ) and acceptor S region germline transcripts (I γ 1-C γ 1, I γ 3-C γ 3, I γ 2b-C γ 2b, and I γ 2a-C γ 2a) by RT-qPCR in activated *Med1*^{F/F}*Mb1*^{Cre/+} and control B cells (Fig. 3 D). We found that the level of I μ -C μ germline transcripts was increased by *Med1* deletion (Fig. 3 D), consistent with the fact that these transcripts accumulate when CSR is not efficient (Thomas-Claudepierre et al., 2013). Interestingly, however, we found that the level of all the acceptor S region germline transcripts was significantly reduced in *Med1*-deficient B cells when compared with control B cells (Fig. 3 D). This is consistent with reduced levels of S α transcripts resulting from Med1 or Med12 knockdown in CH12 cells (Fig. 2 E). We conclude that Med1 is required for robust transcription of the acceptor S regions. This could be explained by two nonmutually exclusive hypotheses. Med1 could be required to initiate transcription at the acceptor S region promoters. Alternatively, Med1 could mediate the long-range interactions between the IgH enhancers and the acceptor S region that are necessary to activate transcription, to bring the acceptor S region to the proximity of S μ and to successfully undergo CSR.

Med1 deficiency results in reduced interactions between E μ and the C γ 3, C γ 1, C γ 2b, and C ϵ genes

To evaluate whether the dynamic three-dimensional changes occurring at the IgH locus during CSR are dependent on the Med1 subunit of the mediator complex, we performed high-resolution 4C-Seq (circular chromosome conformation capture) experiments on resting and activated B cells isolated from the spleens of *Med1*^{F/F}*Mb1*^{Cre/+} and control (*Med1*^{F/F}) mice (Fig. 4 and not depicted). Purified B cells from three different mice per genotype were cultured (or not) for 48 h in the presence of LPS and IL-4 to induce CSR to IgG1 or with LPS alone to induce CSR to IgG3 and IgG2b. As a viewpoint, we used a bait located at the E μ enhancer.

In control resting B cells, we observed a strong interaction between the two IgH enhancers, E μ and 3' RR, that was substantially increased upon activation (Fig. 4 and not depicted). This is consistent with the long-range interactions described in primary B cells and in CH12 B cells (Wuerffel et al., 2007; Pefanis et al., 2015) and with the promoter-enhancer interactome of the IgH locus revealed by Pol II ChIA-PET in primary B cells (Qian et al., 2014). We also observed close-range contacts between E μ and the C μ and C δ genes in resting B cells, which were diminished upon stimulation (Fig. 4 and not depicted). Consistent with the fact that culturing B cells with LPS and IL-4 induces CSR primarily to IgG1, we observed a significant increase in the interaction between E μ and the γ 1 promoter (γ 1p), S γ 1, and the γ 1E enhancer (Fig. 4 and not depicted). Conversely, LPS stimulation led to a significant increase in the interaction between E μ and the γ 2b promoter and S γ 2b, as well as the γ 1E enhancer. Importantly, no interaction was detected between E μ and S γ 1 (Fig. 4 and not depicted), confirming that the formation of loops involving recombining acceptor S regions is stimulation dependent (Wuerffel et al., 2007). The γ 3 promoter and the

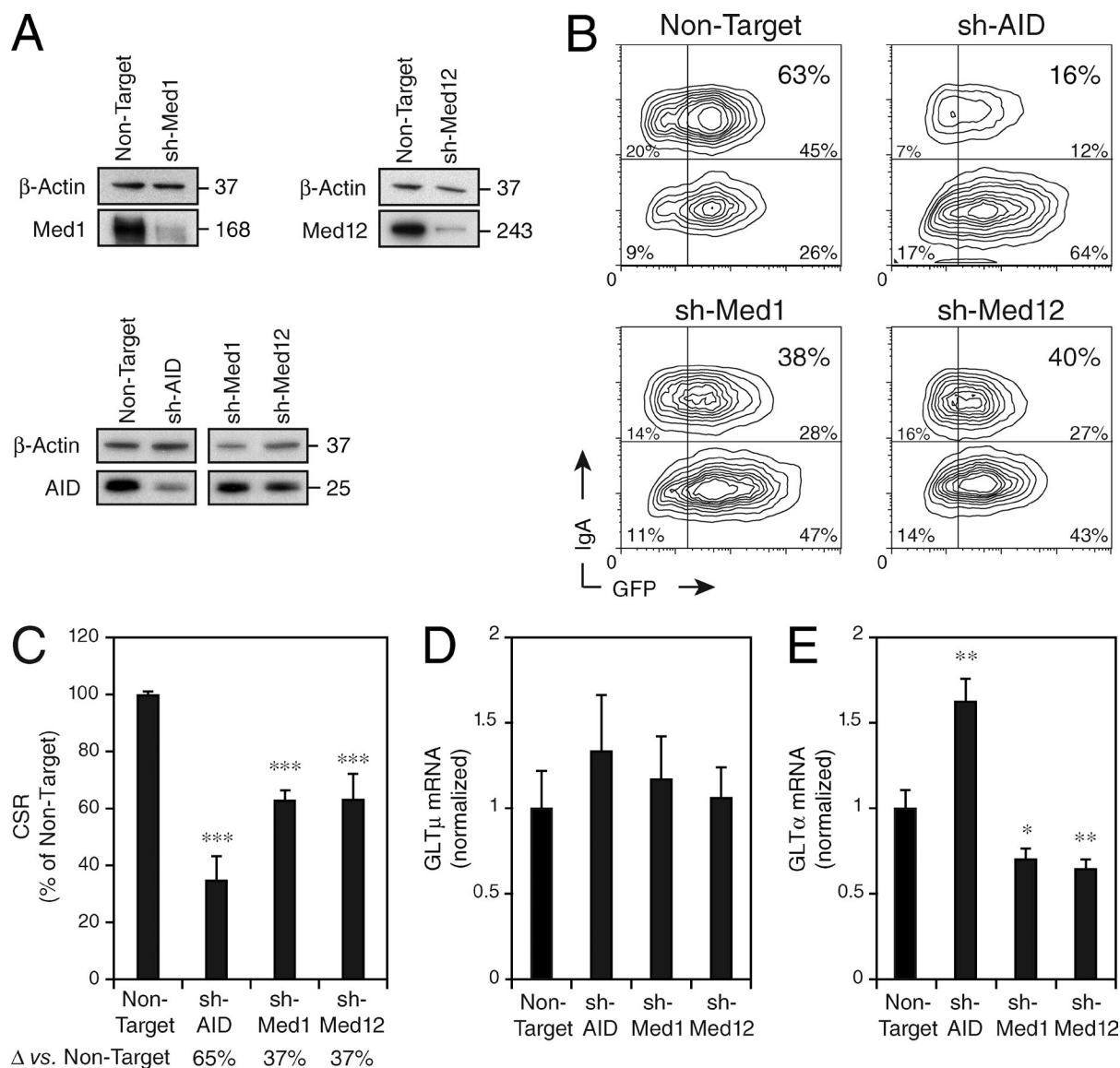


Figure 2. Knockdown of Med1 or Med12 impairs CSR in CH12 cells. (A) Knockdown efficiencies were determined by Western blot on transduced cells stimulated for 48 h and sorted for GFP expression. Western blots for β-actin, Med1, Med12, and AID are shown. Theoretical molecular masses in kilodaltons are indicated. Data are representative of three experiments. (B) IgA surface expression as determined by flow cytometry in stimulated CH12 cells transduced with a lentivirus expressing a GFP reporter and shRNAs specific for AID, Med1, Med12, and a nontarget shRNA negative control. Representative plots (gated on live cells) are shown. Percentage of cells in each quadrant is indicated, and the percentage of IgA⁺ cells among the GFP⁺ population is indicated in the upper right quadrants. (C) Percentage (+SD) of CSR relative to the nontarget shRNA control from four independent experiments by gating on cells expressing GFP. CSR in cells expressing the nontarget shRNA control was set to 100%. The difference in CSR efficiency between nontarget and shRNA knockdown (Δ) is indicated below. Statistical significance versus the nontarget shRNA control (two-tailed Student's *t* test) is indicated. (D and E) RT-qPCR for μ-Cμ (D) and α-Cα (E) germline transcripts in transduced cells stimulated for 48 h and sorted for GFP expression. Transcript cycle threshold values were normalized to hypoxanthine-guanine phosphoribosyltransferase mRNA abundance and are presented relative to the nontarget shRNA negative control (set as 1). Statistical significance versus the nontarget shRNA control (two-tailed Student's *t* test) is indicated. *, *P* ≤ 0.05; **, *P* ≤ 0.01; ***, *P* ≤ 0.0001. Data are representative of three independent experiments.

Sy3 S region appear to be involved in a constitutive interaction with Eμ in resting B cells (Fig. 4 and not depicted). Although this interaction is significantly lost when B cells are stimulated to undergo CSR to IgG1 or IgE but not IgG3 (LPS + IL-4; Fig. 4 and not depicted), it is maintained when

B cells are stimulated to undergo CSR to IgG3 but not IgG1 or IgE (LPS alone; Fig. 4 and not depicted).

Interestingly, the three-dimensional contacts revealed by 4C-Seq in control activated B cells (Fig. 4 and not depicted) perfectly correlate with Med1 and Med12 binding (Fig. 1).

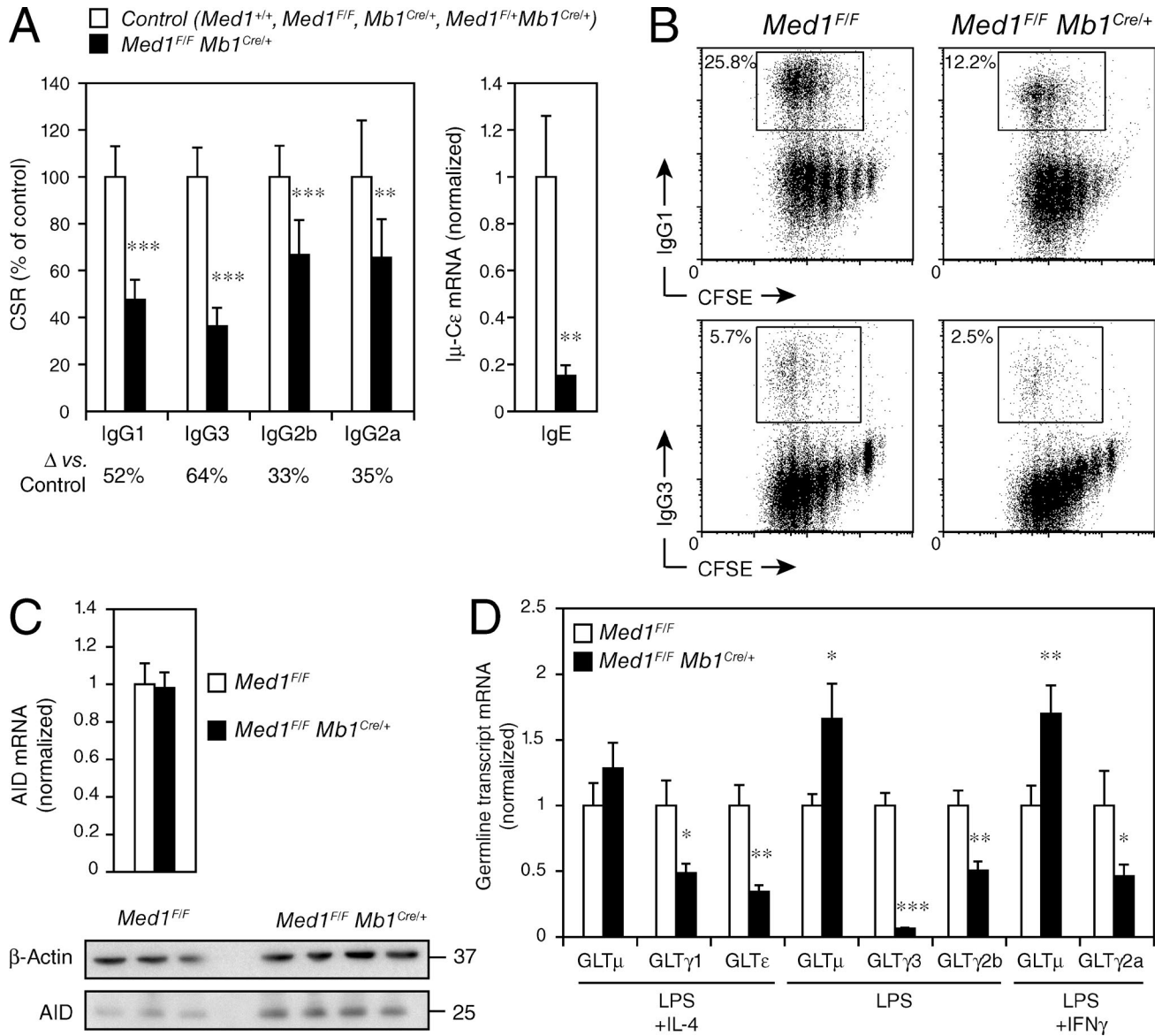


Figure 3. CSR and acceptor S region transcription are compromised by *Med1* deficiency in primary B cells. (A, left) Percentage (+SD) of CSR relative to control cells from three to six independent experiments. The genotypes tested and number of mice were as follows: *Med1^{F/F}Mb1^{Cre/+}* ($n = 37$), *Med1^{+/+}* ($n = 6$), *Mb1^{Cre/+}* ($n = 4$), *Med1^{F/F}* ($n = 28$), *Med1^{F/+}* ($n = 11$), or *Med1^{F/+}Mb1^{Cre/+}* ($n = 16$). No difference between control genotypes (*Med1^{+/+}*, *Mb1^{Cre/+}*, *Med1^{F/F}*, *Med1^{F/+}*, and *Med1^{F/+}Mb1^{Cre/+}*) was observed. CSR in control cells was set to 100%. Statistical analysis was performed using Student's *t* test. **, $P \leq 0.01$; ***, $P \leq 0.0001$. Right: CSR to IgE was evaluated by the levels of I μ -C ϵ post-switch transcripts by RT-qPCR in control and *Med1^{F/F}Mb1^{Cre/+}* B cells cultured for 72 h with LPS + IL-4. Expression is normalized to Ig β and is presented relative to expression in control B cells (set as 1). Mean and SD of triplicate samples are shown. Statistical analysis was performed using two-tailed Student's *t* test. **, $P \leq 0.01$. Data are representative of three experiments with two mice per genotype. (B) Representative example of surface expression of IgG1, IgG3, and CFSE dilution as determined by flow cytometry in *Med1^{F/F}* and *Med1^{F/F}Mb1^{Cre/+}* B cells stimulated for 72 h with LPS + IL-4 or LPS alone. Percentage of switched cells is indicated. (C, top) RT-qPCR analysis for AID mRNA in control and *Med1^{F/F}Mb1^{Cre/+}* B cells cultured for 72 h with LPS + IL-4. Expression is normalized to Ig β and is presented relative to expression in control B cells (set as 1). Mean and SD of triplicate samples are shown. Statistical analysis was performed using two-tailed Student's *t* test. Data are representative of three experiments with two mice per genotype. Bottom: Western blot for β -actin and AID on whole-cell extracts from splenic B cells from *Med1^{F/F}* and *Med1^{F/F}Mb1^{Cre/+}* mice cultured for 72 h with LPS and IL-4. Theoretical molecular masses in kilodaltons are indicated. Data are representative of three independent experiments. (D) RT-qPCR analysis for germline transcripts (GLT) at donor and acceptor S regions in *Med1^{F/F}Mb1^{Cre/+}* and control (*Med1^{F/F}*) B cells cultured for 72 h with LPS alone or with LPS + IL-4 or LPS + IFN- γ . Expression is normalized to Ig β and is presented relative to expression in control B cells, set as 1. Mean and SD of triplicate samples are shown. Statistical analysis was performed using two-tailed Student's *t* test. *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$. Data are representative of three independent experiments with two mice per genotype.

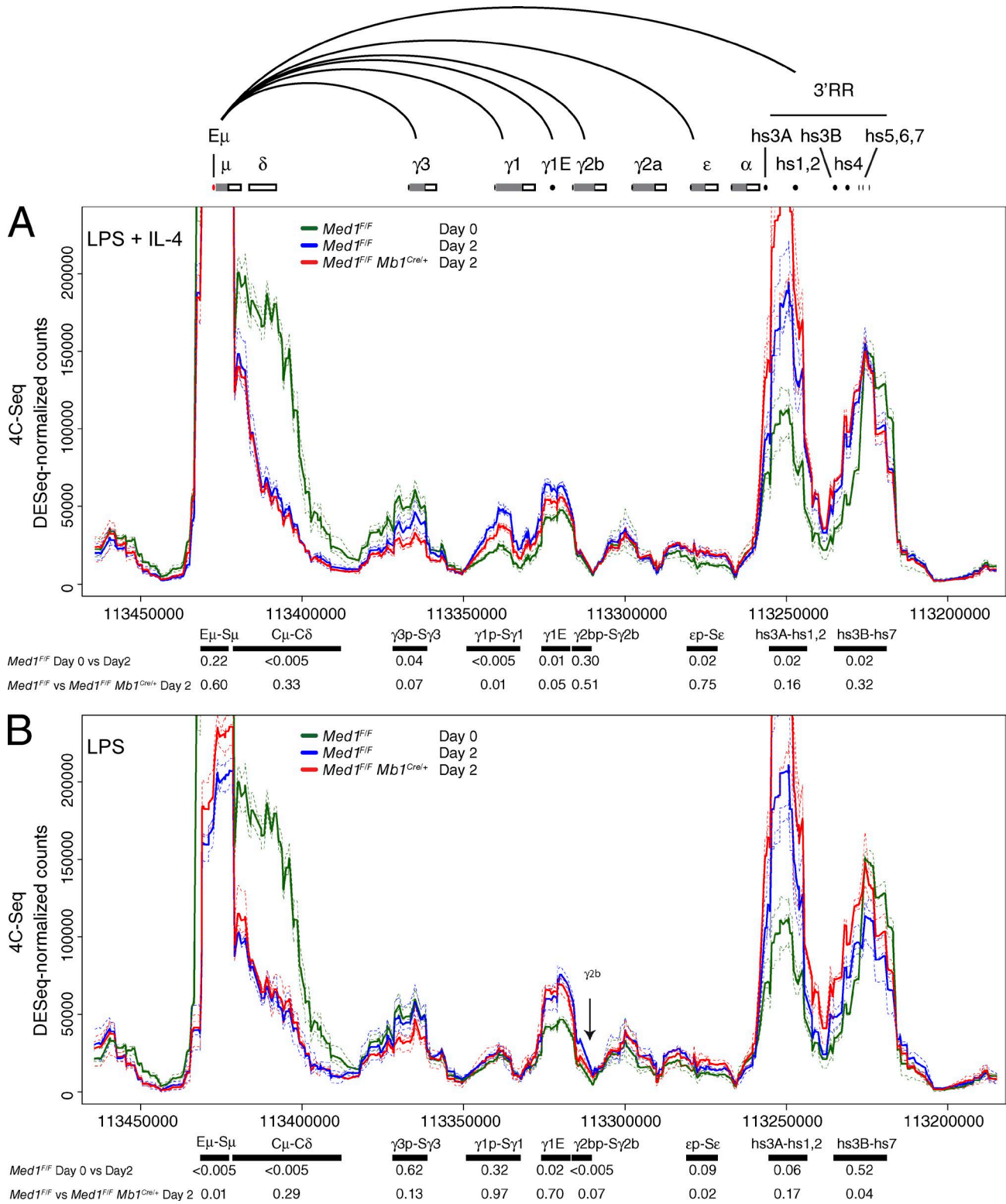


Figure 4. **Med1** deficiency affects IgH chromatin dynamics during CSR. (A and B) High-resolution 4C-Seq was performed using a bait on the Eμ enhancer (red dot). A schematic map of the IgH locus indicates the I exons (black dots), S regions (gray boxes), the constant region exons (white boxes), the Eμ enhancer, the γ1 enhancer (γ1E), and the DNaseI hypersensitive sites (hs) located in the 3' RR. Curved lines indicate long-range interactions. 4C-Seq signal was calculated using 10-kb windows centered on DpnII sites located in the constant region of IgH (chr12: 113175000–113475000, mm10). Full lines represent the mean of three replicates and dashed lines the signal for each replicate. The 4C-Seq analysis was conducted with three biological replicates

In $Med1^{F/F}Mb1^{Cre/+}$ B cells, architectural changes within the IgH locus exhibit a similar profile upon activation (Fig. 4 and not depicted). Nevertheless, the long-range interactions between $E\mu$ and the $\gamma 1p$ - $S\gamma 1$ region and $\gamma 1E$ enhancer were no longer significant (Fig. 4 and not depicted). This suggests a role for Med1 in facilitating these chromosomal interactions that occur after activation.

Across the entire IgH locus, no significant differences could be detected when comparing activated $Med1^{F/F}Mb1^{Cre/+}$ with control B cells (Fig. 4). However, when we focused our analysis on the regions whose interactions with $E\mu$ change upon activation and whose transcription is induced in a stimulation-dependent manner (Fig. 4), we detected significant differences between $Med1^{F/F}Mb1^{Cre/+}$ and control B cells (Fig. 4). Notably, the interaction between $E\mu$ and the $\gamma 1p$ - $S\gamma 1$ region and $\gamma 1E$ enhancer was significantly reduced ($P = 0.01$ and $P = 0.05$, respectively) in activated (LPS + IL-4) $Med1^{F/F}Mb1^{Cre/+}$ B versus control B cells (Fig. 4 A). This shows that the $C\gamma 1$ gene is not efficiently brought into close proximity with $S\mu$ in $Med1$ -deficient B cells. Similarly, $E\mu$ - $\gamma 2bp$ - $S\gamma 2b$ interactions induced by LPS tend to be less frequent in $Med1$ -deficient B cells (Fig. 4 B). The lower robustness of CSR induction by LPS alone (when compared with LPS + IL-4), in which only a small percentage of cells underwent CSR (Fig. 3 B), could explain why these interactions are not as frequent in control cells and why the difference between $Med1$ -deficient and control B cells is not more pronounced.

Altogether, the decrease in interactions observed in $Med1$ -deficient B cells relative to control B cells (Fig. 4) is consistent with the reduction in acceptor S region transcription (Fig. 3 D) and the defect in CSR observed (Fig. 3, A and B). We conclude that the long-range contacts occurring in resting B cells between both of the IgH enhancers are robust and remain intact upon $Med1$ deletion. In contrast, the dynamic three-dimensional changes occurring at the IgH locus upon activation between $E\mu$ and the $C\gamma 3$, $C\gamma 1$, $C\gamma 2b$, and $C\epsilon$ genes are facilitated by the Med1 subunit of the mediator complex. It is impossible, however, to determine whether the reduced frequency of interaction between $E\mu$ and the C genes observed in $Med1$ -deficient B cells is the cause or the consequence of reduced transcription at the acceptor S regions.

$Med1$ deficiency results in a 50% reduction in the efficiency of CSR, and it is possible that those B cells, which successfully underwent CSR in the absence of Med1, managed to properly synapse and transcribe the donor and acceptor S regions by a mediator-independent mechanism. Alternatively, as the Med12 subunit can be recruited to the

IgH locus in the absence of Med1 (Fig. 1, A and B) and mediator complexes lacking the Med1 subunit can exist (Borggreffe and Yue, 2011; Allen and Taatjes, 2015), it is possible that the $Med1$ -deficient B cells that managed to undergo CSR were able to assemble and recruit a mediator complex independently of the Med1 subunit. Our results implicate the mediator complex in the mechanism of CSR and are consistent with a model in which mediator, possibly together with cohesin, facilitates the transcriptional activation of S regions and their long-range contacts with the IgH locus enhancers during CSR.

MATERIALS AND METHODS

Mice. $Mb1^{Cre/+}$ mice (Hobeika et al., 2006) were obtained from M. Reth (Max Planck Institute of Immunobiology and Epigenetics, Freiburg, Germany). $Med1^{F/F}$ (Jia et al., 2004) and $Mb1^{Cre/+}$ mice were on a B6/129 mixed background and bred and maintained under specific pathogen-free conditions. 8–12-wk-old mice were used in all experiments. All animal work was performed under protocols approved by the Direction des Services Vétérinaires du Bas-Rhin (authorization number 67-343).

Lentiviral transduction. The lentiviral vectors (pLKO.1 and pLKO.1-puro-CMV-TurboGFP) expressing shRNAs specific for AID (TRCN0000112031), Med1 (TRCN0000099576), Med12 (TRCN0000096466), or a nontarget control (SHC002) were obtained from Sigma-Aldrich. CH12 cells were infected with viral particles produced in Lenti-X 293T cells as described previously (Thomas-Claudepierre et al., 2013).

Cell culture and flow cytometry. Resting splenic B cells were isolated using CD43 microbeads (Miltenyi Biotec), stained with 5 μ M CFSE (Invitrogen), and cultured for 72 h with 50 μ g/ml LPS (Sigma-Aldrich) to switch to IgG3 and IgG2b, LPS and IL-4 (5 ng/ml; PeproTech) to switch to IgG1, and LPS and IFN- γ (100 ng/ml; PeproTech) to switch to IgG2a. CSR was assayed by flow cytometry as described previously (Jeevan-Raj et al., 2011). CH12 cells were transduced with shRNA-expressing lentiviral vectors and were cultured with IL-4, TGF- β , and anti-CD40 antibody to switch to IgA as described previously (Thomas-Claudepierre et al., 2013). After 72 h, enhanced GFP expression and surface IgA was analyzed by flow cytometry.

Western blot analysis. Proteins were fractionated by SDS-PAGE on 4–12% gradient gels (Invitrogen), transferred to Immobilon polyvinylidene difluoride membranes (EMD

per time point (day 0 and day 2), genotype ($Med1^{F/F}$ and $Med1^{F/F}Mb1^{Cre/+}$), and stimulation: resting B cells and LPS + IL-4-stimulated cells (A) and resting B cells and LPS-stimulated cells (B). We focused our analysis on the RRs and on the regions whose transcription is specifically induced upon stimulation with LPS or LPS + IL-4 (black bars). In this focused analysis, a Welch's t test (two-sample and two-sided) was used on the mean score for all windows centered within the chosen regions to assess significant differences, taking into account the three replicates for each sample. P-values are indicated. The $\gamma 2bp$ - $S\gamma 2b$ region is indicated with an arrow.

Millipore), and analyzed by Western blot. See Table S2 for antibodies used in this study.

ChIP-qPCR. Analytical scale ChIP was performed on chromatin prepared from 10^7 (resting or activated) splenic B cells isolated from a pool of two to three mice as described previously (Yamane et al., 2011). qPCR was performed at several locations across the IgH locus using primers listed in Table S3. Results are expressed as percent input and represent the mean of three biological replicates. Error bars represent the corresponding SD.

ChIP-Seq analysis. Sequence reads (SRR975483 [Wang et al., 2014] and SRR620145 [Whyte et al., 2013; Predeus et al., 2014]) were mapped to reference genome mm10/GRCm38 using Bowtie v1.12.8. Peak calling was performed using MACS (Zhang et al., 2008) with default parameters.

RT-qPCR. RNA and cDNA were prepared using standard techniques. qPCR was performed in triplicates using the Universal Probe Library system (Roche) or SyberGreen (QIA GEN) and a LightCycler 480 (Roche). Transcript quantities were calculated relative to standard curves and normalized to β -actin, Ig β , or hypoxanthine-guanine phosphoribosyltransferase (HPRT) mRNA. See Table S3 for primers and probes used in this study.

High-resolution 4C-Seq. 4C-Seq was performed as described previously (Rocha et al., 2012). The primary restriction enzyme used was DpnII and secondary restriction enzyme Csp6I. Primers for the E μ bait were 5'-TCTGTCCTAAAG GCTCTGAGA-3' and 5'-GAACACAGAAGTATGTGT ATGGA-3'. The 18 samples (three biological replicates, two genotypes, and three stimulation conditions) were sequenced on a HiSeq2500 rapid run using 50 cycles. Mapping was done using Bowtie, allowing for zero mismatches against a reduced genome containing all unique 24 bp surrounding every DpnII site in the genome. Between 2 and 4 million mapped reads were obtained for all samples, and at least half of the reads were mapped to chromosome 12. DESeq 1.6.0 was used to normalize raw read count in 10-kb windows centered on each DpnII site in the constant region of IgH (chr12: 113175000–113475000, mm10; Anders and Huber, 2010). The following parameters were used with the estimate Dispersion function: method = pooled, sharingMode = maximum. The nbinomTest function was used to determine which windows have a significantly different 4C-Seq signal. An adjusted p-value of 0.01 after Benjamini-Hochberg false discovery rate correction was used as a cutoff. For the analysis of specific regions, a Welch's *t* test (two-sample and two-sided) on the mean score for all windows centered within the chosen regions was used to assess significant differences, taking into account the three replicates for each sample. The coordinates used were as follows: E μ -S μ , chr12:113423031–113428514; C μ -C δ , chr12:113384000–113420500; γ 3p-S γ 3,

chr12:113361533–113371500; γ 1p-S γ 1, chr12:113332600–113349000; γ 1E, chr12:113319346–113325934; γ 2bp-S γ 2b, chr12:113310111–113315500; ϵ p-Se, chr12:113273549–113280029; hs1,2-hs3A, chr12:113242500–113255414; and hs3B/hs7, chr12:113220000–113233375.

Accession number. 4C-Seq data on resting and activated, control, and *Med1*-deficient B cells were submitted to the Gene Expression Omnibus (accession number GSE62969).

Online supplemental material. Table S1 includes statistics for ChIP-qPCR (Fig. 1). Table S2 lists antibodies used in this study. Table S3 lists primers and probes used in this study. Online supplemental material is available at <http://www.jem.org/cgi/content/full/jem.20141967/DC1>.

ACKNOWLEDGMENTS

We thank members of the Reina-San-Martin laboratory for discussions, E. Soutoglou, T. Sexton, and A. Ramiro for comments on the manuscript, C. Ebel for assistance with cell sorting, M. Gendron for animal care, M. Reth for *Mb1^{Cre}* mice, and the New York University Medical Center genome technology center for their contributions to this work.

This work was supported by grants to B. Reina-San-Martin from the Agence Nationale pour la Recherche (ANR-Bland) and the Association for Research on Cancer Foundation (Programme ARC). A.-S. Thomas-Claudepierre was supported by the Ministère de l'Enseignement Supérieur et de la Recherche and the Association for Research on Cancer Foundation. P.P. Rocha is supported by an American Society of Hematology fellowship. J.A. Skok is supported by National Institutes of Health grant R01 GM086852. J.A. Skok and R. Bonneau are supported by National Institutes of Health grant R01GM112192. This study was supported by the grant ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the program Investissements d'Avenir labeled ANR-10-IDEX-0002-02.

The authors declare no competing financial interests.

Submitted: 17 October 2014

Accepted: 15 January 2016

REFERENCES

- Allen, B.L., and D.J. Taatjes. 2015. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* 16:155–166. <http://dx.doi.org/10.1038/nrm3951>
- Anders, S., and W. Huber. 2010. Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. <http://dx.doi.org/10.1186/gb-2010-11-10-r106>
- Basu, U., F.L. Meng, C. Keim, V. Grinstein, E. Pefanis, J. Eccleston, T. Zhang, D. Myers, C.R. Wasserman, D.R. Wesemann, et al. 2011. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell.* 144:353–363. <http://dx.doi.org/10.1016/j.cell.2011.01.001>
- Borggrefe, T., and X. Yue. 2011. Interactions between subunits of the Mediator complex with gene-specific transcription factors. *Semin. Cell Dev. Biol.* 22:759–768. <http://dx.doi.org/10.1016/j.semcdb.2011.07.022>
- Chaudhuri, J., U. Basu, A. Zarrin, C. Yan, S. Franco, T. Perlot, B. Vuong, J. Wang, R. T. Phan, A. Datta, et al. 2007. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv. Immunol.* 94:157–214. [http://dx.doi.org/10.1016/S0065-2776\(06\)94006-1](http://dx.doi.org/10.1016/S0065-2776(06)94006-1)
- Conaway, R.C., and J.W. Conaway. 2011. Origins and activity of the Mediator complex. *Semin. Cell Dev. Biol.* 22:729–734. <http://dx.doi.org/10.1016/j.semcdb.2011.07.021>
- Feldman, S., I. Achour, R. Wuerffel, S. Kumar, T. Gerasimova, R. Sen, and A.L. Kenter. 2015. Constraints contributed by chromatin looping limit

- recombination targeting during Ig class switch recombination. *J. Immunol.* 194:2380–2389. <http://dx.doi.org/10.4049/jimmunol.1401170>
- Hobeika, E., S. Thiemann, B. Storch, H. Jumaa, P.J. Nielsen, R. Pelanda, and M. Reth. 2006. Testing gene function early in the B cell lineage in mb1-cre mice. *Proc. Natl. Acad. Sci. USA.* 103:13789–13794. <http://dx.doi.org/10.1073/pnas.0605944103>
- Ito, M., C.X. Yuan, H.J. Okano, R.B. Darnell, and R.G. Roeder. 2000. Involvement of the TRAP220 component of the TRAP/SMCC coactivator complex in embryonic development and thyroid hormone action. *Mol. Cell.* 5:683–693. [http://dx.doi.org/10.1016/S1097-2765\(00\)80247-6](http://dx.doi.org/10.1016/S1097-2765(00)80247-6)
- Jeevan-Raj, B.P., I. Robert, V. Heyer, A. Page, J.H. Wang, F. Cammas, F.W. Alt, R. Losson, and B. Reina-San-Martin. 2011. Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *J. Exp. Med.* 208:1649–1660. <http://dx.doi.org/10.1084/jem.20110118>
- Jia, Y., C. Qi, P. Kashireddi, S. Surapureddi, Y.J. Zhu, M.S. Rao, D. Le Roith, P. Chambon, F.J. Gonzalez, and J.K. Reddy. 2004. Transcription coactivator PBP, the peroxisome proliferator-activated receptor (PPAR)-binding protein, is required for PPAR α -regulated gene expression in liver. *J. Biol. Chem.* 279:24427–24434. <http://dx.doi.org/10.1074/jbc.M402391200>
- Kagey, M.H., J.J. Newman, S. Bilodeau, Y. Zhan, D.A. Orlando, N.L. van Berkum, C.C. Ebmeier, J. Goossens, P.B. Rahl, S.S. Levine, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature.* 467:430–435. <http://dx.doi.org/10.1038/nature09380>
- Kenter, A.L., S. Feldman, R. Wuerrfel, I. Achour, L. Wang, and S. Kumar. 2012. Three-dimensional architecture of the IgH locus facilitates class switch recombination. *Ann. N.Y. Acad. Sci.* 1267:86–94. <http://dx.doi.org/10.1111/j.1749-6632.2012.06604.x>
- Malik, S., and R.G. Roeder. 2010. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat. Rev. Genet.* 11:761–772. <http://dx.doi.org/10.1038/nrg2901>
- Medvedovic, J., A. Ebert, H. Tagoh, I.M. Tamir, T.A. Schwickert, M. Novatchkova, Q. Sun, P.J. Huis In 't Veld, C. Guo, H.S. Yoon, et al. 2013. Flexible long-range loops in the VH gene region of the Igh locus facilitate the generation of a diverse antibody repertoire. *Immunity.* 39:229–244. <http://dx.doi.org/10.1016/j.immuni.2013.08.011>
- Pavri, R., and M.C. Nussenzweig. 2011. AID targeting in antibody diversity. *Adv. Immunol.* 110:1–26. <http://dx.doi.org/10.1016/B978-0-12-387663-8.00005-3>
- Pefanis, E., J. Wang, G. Rothschild, J. Lim, D. Kazadi, J. Sun, A. Federation, J. Chao, O. Elliott, Z.P. Liu, et al. 2015. RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell.* 161:774–789. <http://dx.doi.org/10.1016/j.cell.2015.04.034>
- Petersen-Mahrt, S.K., R.S. Harris, and M.S. Neuberger. 2002. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature.* 418:99–104. <http://dx.doi.org/10.1038/nature00862>
- Phillips-Cremins, J.E., M.E. Sauria, A. Sanyal, T.I. Gerasimova, B.R. Lajoie, J.S. Bell, C.T. Ong, T.A. Hookway, C. Guo, Y. Sun, et al. 2013. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell.* 153:1281–1295. <http://dx.doi.org/10.1016/j.cell.2013.04.053>
- Predeus, A.V., S. Gopalakrishnan, Y. Huang, J. Tang, A.J. Feeney, E.M. Oltz, and M.N. Artyomov. 2014. Targeted chromatin profiling reveals novel enhancers in Ig H and Ig L chain Loci. *J. Immunol.* 192:1064–1070. <http://dx.doi.org/10.4049/jimmunol.1302800>
- Qian, J., Q. Wang, M. Dose, N. Pruett, K.R. Kieffer-Kwon, W. Resch, G. Liang, Z. Tang, E. Mathé, C. Benner, et al. 2014. B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. *Cell.* 159:1524–1537. <http://dx.doi.org/10.1016/j.cell.2014.11.013>
- Rocha, P.P., M. Micsinai, J.R. Kim, S.L. Hewitt, P.P. Souza, T. Trimarchi, F. Strino, F. Parisi, Y. Kluger, and J.A. Skok. 2012. Close proximity to Igh is a contributing factor to AID-mediated translocations. *Mol. Cell.* 47:873–885. <http://dx.doi.org/10.1016/j.molcel.2012.06.036>
- Stumpf, M., X. Yue, S. Schmitz, H. Luche, J.K. Reddy, and T. Borggrefe. 2010. Specific erythroid-lineage defect in mice conditionally deficient for Mediator subunit Med1. *Proc. Natl. Acad. Sci. USA.* 107:21541–21546. <http://dx.doi.org/10.1073/pnas.1005794107>
- Thomas-Claudepierre, A.S., E. Schiavo, V. Heyer, M. Fournier, A. Page, I. Robert, and B. Reina-San-Martin. 2013. The cohesin complex regulates immunoglobulin class switch recombination. *J. Exp. Med.* 210:2495–2502. <http://dx.doi.org/10.1084/jem.20130166>
- Wang, Q., T. Oliveira, M. Jankovic, I.T. Silva, O. Hakim, K. Yao, A. Gazumyan, C.T. Mayer, R. Pavri, R. Casellas, et al. 2014. Epigenetic targeting of activation-induced cytidine deaminase. *Proc. Natl. Acad. Sci. USA.* 111:18667–18672. <http://dx.doi.org/10.1073/pnas.1420575111>
- Whyte, W.A., D.A. Orlando, D. Hnisz, B.J. Abraham, C.Y. Lin, M.H. Kagey, P.B. Rahl, T.I. Lee, and R.A. Young. 2013. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell.* 153:307–319. <http://dx.doi.org/10.1016/j.cell.2013.03.035>
- Wuerrfel, R., L. Wang, F. Grigera, J. Manis, E. Selsing, T. Perlot, F.W. Alt, M. Cogne, E. Pinaud, and A.L. Kenter. 2007. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity.* 27:711–722. <http://dx.doi.org/10.1016/j.immuni.2007.09.007>
- Yamane, A., W. Resch, N. Kuo, S. Kuchen, Z. Li, H.W. Sun, D.F. Robbani, K. McBride, M.C. Nussenzweig, and R. Casellas. 2011. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat. Immunol.* 12:62–69. <http://dx.doi.org/10.1038/ni.1964>
- Yue, X., A. Izcue, and T. Borggrefe. 2011. Essential role of Mediator subunit Med1 in invariant natural killer T-cell development. *Proc. Natl. Acad. Sci. USA.* 108:17105–17110. <http://dx.doi.org/10.1073/pnas.1109095108>
- Zhang, Y., T. Liu, C.A. Meyer, J. Eeckhoutte, D.S. Johnson, B.E. Bernstein, C. Nusbaum, R.M. Myers, M. Brown, W. Li, and X.S. Liu. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137. <http://dx.doi.org/10.1186/gb-2008-9-9-r137>
- Zhu, Y., C. Qi, Y. Jia, J.S. Nye, M.S. Rao, and J.K. Reddy. 2000. Deletion of PBP/PPARBP, the gene for nuclear receptor coactivator peroxisome proliferator-activated receptor-binding protein, results in embryonic lethality. *J. Biol. Chem.* 275:14779–14782. <http://dx.doi.org/10.1074/jbc.C000121200>

Fam72a enforces error-prone DNA repair during antibody diversification

<https://doi.org/10.1038/s41586-021-04093-y>

Received: 17 December 2020

Accepted: 4 October 2021

Published online: 24 November 2021

 Check for updates

Mélanie Rogier^{1,2,3,4,12}, Jacques Moritz^{1,2,3,4,12}, Isabelle Robert^{1,2,3,4}, Chloé Lescale⁵, Vincent Heyer^{1,2,3,4}, Arthur Abello^{1,2,3,4}, Ophélie Martin⁶, Katia Capitani^{7,8}, Morgane Thomas⁹, Anne-Sophie Thomas-Claudepierre^{1,2,3,4}, Brice Laffleur¹⁰, Florence Jouan¹⁰, Eric Pinaud⁹, Karin Tarte¹⁰, Michel Cogne^{9,10}, Silvestro G. Conticello^{7,11}, Evi Soutoglou⁶, Ludovic Deriano⁵ & Bernardo Reina-San-Martin^{12,3,4}✉

Efficient humoral responses rely on DNA damage, mutagenesis and error-prone DNA repair. Diversification of B cell receptors through somatic hypermutation and class-switch recombination are initiated by cytidine deamination in DNA mediated by activation-induced cytidine deaminase (AID)¹ and by the subsequent excision of the resulting uracils by uracil DNA glycosylase (UNG) and by mismatch repair proteins^{1–3}. Although uracils arising in DNA are accurately repaired^{1–4}, how these pathways are co-opted to generate mutations and double-strand DNA breaks in the context of somatic hypermutation and class-switch recombination is unknown^{1–3}. Here we performed a genome-wide CRISPR–Cas9 knockout screen for genes involved in class-switch recombination and identified FAM72A, a protein that interacts with the nuclear isoform of UNG (UNG2)⁵ and is overexpressed in several cancers⁵. We show that the FAM72A–UNG2 interaction controls the levels of UNG2 and that class-switch recombination is defective in *Fam72a*^{−/−} B cells due to the upregulation of UNG2. Moreover, we show that somatic hypermutation is reduced in *Fam72a*^{−/−} B cells and that its pattern is skewed upon upregulation of UNG2. Our results are consistent with a model in which FAM72A interacts with UNG2 to control its physiological level by triggering its degradation, regulating the level of uracil excision and thus the balance between error-prone and error-free DNA repair. Our findings have potential implications for tumorigenesis, as reduced levels of UNG2 mediated by overexpression of *Fam72a* would shift the balance towards mutagenic DNA repair, rendering cells more prone to acquire mutations.

During immune responses, mammalian B cells undergo somatic hypermutation (SHM) and class-switch recombination (CSR) to diversify the B cell receptor repertoire^{1–3}. SHM modifies antibody affinity by mutating the variable region of both heavy (IgH) and light (IgL) chains^{1–3}. CSR modulates antibody effector functions by replacing the isotype expressed from IgM/IgD to IgG, IgA or IgE, through a recombination event taking place at the *Igh* locus^{1–3}. SHM and CSR are triggered by activation-induced cytidine deaminase (AID), which deaminates cytosines into uracils in DNA^{1–3}. AID-generated U:G mismatches are processed mainly by the base-excision repair enzyme UNG^{3,6} but also by mismatch repair⁷ proteins to introduce mutations or double-strand DNA breaks (DSBs)^{1–3}. During SHM, DNA replication over U:G mismatches generates transition mutations at G:C base pairs. The excision of uracil by UNG, followed by replication over abasic sites introduces

both transition and transversion mutations at C:G base pairs. In addition, single-stranded DNA surrounding the U:G mismatch or abasic site can be excised by both mismatch repair and base-excision repair to introduce transition and transversion mutations at A:T base pairs. During CSR, cytidine deamination and DNA cleavage on opposite strands generate DSBs^{1–3}. While uracils in DNA are usually accurately repaired^{1–3}, the mechanisms by which AID-induced uracils are instead processed in an error-prone way during SHM and CSR are largely unknown^{1–3}.

To gain insight into the molecular mechanisms of CSR, we conducted a genome-wide CRISPR–Cas9 knockout screen (Fig. 1a) using the IgM⁺ murine B cell line CH12 (ref. ⁸), which we engineered to stably express Cas9 (CH12^{Cas9}) Extended Data Fig. 1a) and the Brie⁹ guide RNA (gRNA) library (pMX-mBrie) (Extended Data Fig. 1b). CH12^{Cas9} cells were transduced, selected with puromycin and then stimulated to undergo CSR for

¹Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France. ²Institut National de la Santé et de la Recherche Médicale (INSERM), U1258, Illkirch, France. ³Centre National de la Recherche Scientifique (CNRS), UMR7104, Illkirch, France. ⁴Université de Strasbourg, Illkirch, France. ⁵Genome Integrity, Immunity and Cancer Unit, Equipe Labellisée Ligue Contre Le Cancer, INSERM U1223, Institut Pasteur, Paris, France. ⁶Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Brighton, UK. ⁷Core Research Laboratory, ISPRO, Firenze, Italy. ⁸Department of Medical Biotechnologies, University of Siena, Siena, Italy. ⁹Centre National de la Recherche Scientifique (CNRS), UMR7276, Institut National de la Santé et de la Recherche Médicale (INSERM), UMR1262-Contrôle de la Réponse Immune B et Lymphoproliférations, Université de Limoges, Limoges, France. ¹⁰Institut national de la santé et de la recherche médicale (INSERM), UMR1236, Université Rennes 1, Etablissement Français du Sang Bretagne, Rennes, France. ¹¹Institute of Clinical Physiology, National Research Council, Pisa, Italy. ¹²These authors contributed equally: Mélanie Rogier, Jacques Moritz. ✉e-mail: reinab@igbmc.fr

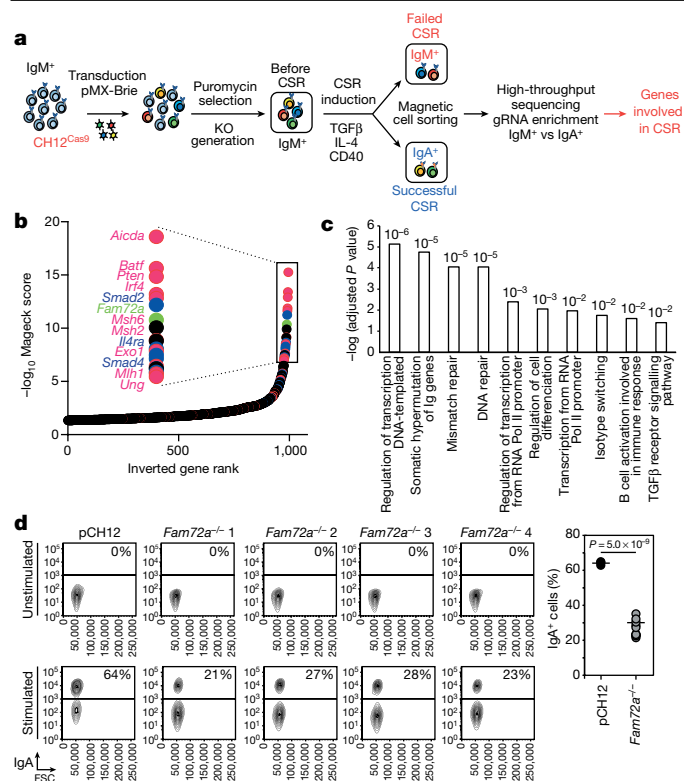


Fig. 1 | A genome-wide CRISPR-Cas9 knockout screen for genes involved in CSR identifies *Fam72a*. **a**, Schematic overview. KO, knockout. **b**, Gene ranking ($-\log_{10}$ Magesk score), plotted against the logarithmic value to the base 2 of the ratio of IgM over IgA read counts (LFC). Genes known to be involved in CSR are depicted in red; TGF β , IL-4, CD40 and B cell receptor signalling pathways are depicted in blue; and genes with no known function in CSR are in black. **c**, Gene ontology analysis of genes identified in the screen. *P* values are indicated. RNA Pol II, RNA polymerase II. **d**, FACS analysis of IgA expression (versus forward scatter (FSC)) in wild-type (pCH12) and four independent *Fam72a*^{-/-} CH12 cell clones stimulated or not with TGF β , anti-CD40 antibody and IL-4. The percentage of IgA⁺ cells from three independent experiments is shown on the right (one-tailed Student's *t*-test). The horizontal line denotes mean values.

72 h. Cells that failed to undergo CSR (which remained IgM⁺) or that succeed (which became IgA⁺) were sorted (Fig. 1a, Extended Data Fig. 1c). Genes, whose knockout affect the efficiency of CSR, were identified by scoring the enrichment of gRNA in the IgM⁺ versus IgA⁺ populations through deep sequencing¹⁰ (Fig. 1a). The screen identified 654 candidate genes (Fig. 1b, Supplementary Table 1) and was successful as it identified genes known to be required for CSR (*Aicda* (which encodes AID), *Ung* and *Trp53bp1*), together with transcription factors that drive AID expression (*Batf*, *Pten* and *Irf4*), non-homologous end-joining factors (*Xrcc5* and *Fam35a*), mismatch repair proteins (*Msh2*, *Msh6*, *Pms1*, *Pms2* and *Mlh1*), members of the TGF β , IL-4 and CD40 signalling pathways (*Tgfb1*, *Tgfb2*, *Il4ra*, *Il2rg* and *Cd40*) and other genes known to be implicated in CSR (*Mediator*, *Trim28* and *Zmynd8*, among other genes). Consistent with this, gene ontology analysis revealed a significant enrichment in pathways that are relevant for CSR and SHM (Fig. 1c). A number of genes with no known function in CSR were identified, including *Fam72a*.

We focused our analysis on *Fam72a*, whose expression is enhanced in B cells undergoing CSR (Extended Data Fig. 1d). *Fam72a* encodes a protein that binds to the nuclear isoform of UNG (UNG2) in a tryptophan 125-dependent manner⁵ (Extended Data Fig. 2a, Supplementary Fig. 1) and is overexpressed in multiple cancers⁵. To determine whether *Fam72a* is required for CSR, we generated *Fam72a*^{-/-} CH12 cells (Extended Data Fig. 2b, c). We found that the efficiency of CSR was

reduced by 50–60% in *Fam72a*^{-/-} CH12 clones (Fig. 1d). The defect in CSR was not due to impaired expression of AID (Extended Data Fig. 2d), defects in proliferation (Extended Data Fig. 2e) or switch region transcription (Extended Data Fig. 2f). We concluded that *Fam72a* is required for efficient CSR in CH12 cells.

To investigate the physiological role of *Fam72a* in CSR and SHM in vivo, we obtained a *Fam72a*^{-/-} mouse model¹¹ (Extended Data Fig. 3a). *Fam72a*^{-/-} mice displayed normal development of B cells (Fig. 2a, b, Extended Data Fig. 3b, c), suggesting that FAM72A is dispensable for V(D)J recombination.

To determine whether *Fam72a*^{-/-} B cells display a defect in CSR, we purified resting splenic B cells and induced CSR to different isotypes. *Fam72a*^{-/-} B cells displayed a 40–70% reduction in CSR (Fig. 2c, d), which was not due to defects in proliferation (Fig. 2c), expression of AID (Fig. 2e) or switch region transcription (Extended Data Fig. 3d). We concluded that deficiency in *Fam72a* results in a B cell-intrinsic CSR defect in mouse primary B cells.

To investigate the functional relevance of the FAM72A-UNG2 interaction, we re-expressed FAM72A and the FAM72A(W125R) or FAM72A(W125A) UNG2-binding defective mutants⁵ (Extended Data Fig. 2a) in *Fam72a*^{-/-} CH12 B cells. Re-expression of FAM72A (Extended Data Fig. 4a) rescued CSR (Fig. 3a, c), whereas FAM72A(W125R) or FAM72A(W125A) did not (Fig. 3a, c). The residual CSR activity in *Fam72a*^{-/-} B cells was dependent on the catalytic activity of UNG (Fig. 3c, Extended Data Fig. 4b). We concluded that the CSR defect observed is due to the lack of FAM72A, that the FAM72A-UNG2 interaction is of functional relevance for CSR and that the residual CSR activity is dependent on UNG.

The *Ung* gene encodes a mitochondrial (UNG1) and a nuclear (UNG2) isoform (Extended Data Fig. 4c), both of which are able to

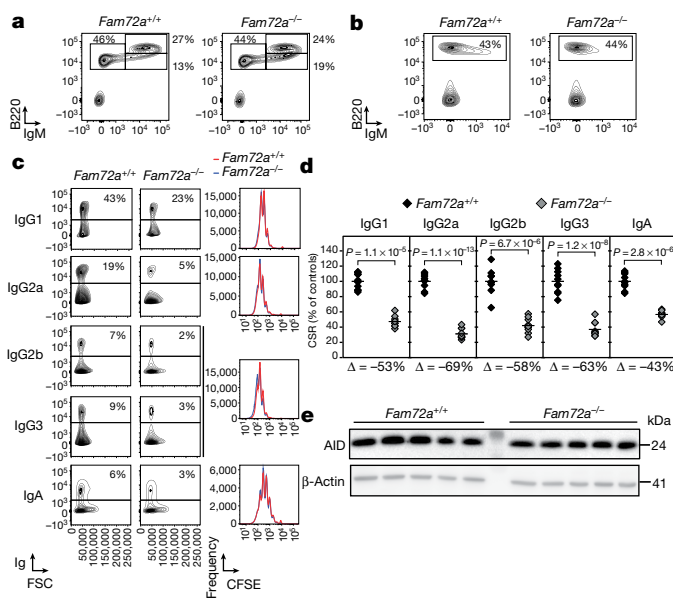


Fig. 2 | CSR is defective in B cells from *Fam72a*^{-/-} mice. **a**, **b**, FACS analysis of bone marrow (**a**) and spleen (**b**) cells from *Fam72a*^{+/+} and *Fam72a*^{-/-} mice. Staining antibodies and the percentage of cells within gates are indicated. **c**, FACS analysis of Ig expression in *Fam72a*^{+/+} and *Fam72a*^{-/-} stimulated splenic B cells. The percentage of Ig-expressing cells is indicated. Representative plots from two experiments with five mice of each genotype are shown. Carboxyfluorescein succinimidyl ester (CFSE) dilution analysis is shown on the right. **d**, Percentage of CSR in *Fam72a*^{-/-} B cells relative to *Fam72a*^{+/+} B cells for the different isotypes from ten independent mice. The horizontal line denotes mean values. The percentage difference (Δ) is indicated below (one-tailed Student's *t*-test). **e**, Western blot analysis for AID and β -actin in *Fam72a*^{+/+} and *Fam72a*^{-/-} stimulated splenic B cells. For gel source data, see Supplementary Fig. 1.

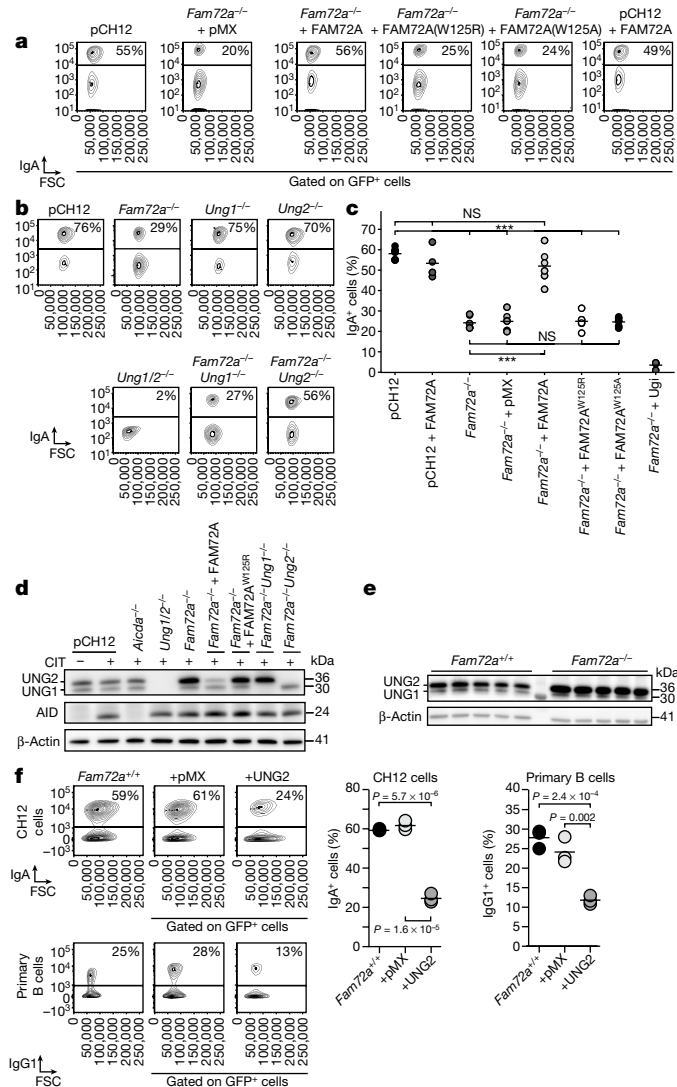


Fig. 3 | The CSR defect in *Fam72a*^{-/-} B cells is due to the specific upregulation of UNG2. **a**, FACS analysis of IgA expression in stimulated *Fam72a*^{-/-} CH12 cells transduced with an empty retrovirus (pMX) or expressing FAM72A, FAM72A(W125A) or FAM72A(W125R). The percentage of IgA-expressing cells is indicated. The parental cell line (pCH12) was included as a control. **b**, FACS analysis of IgA expression in stimulated CH12 cells with the indicated genotype. The percentage of IgA-expressing cells is indicated. **c**, The percentage of IgA⁺ cells from three independent experiments from **a** and Extended Data Fig. 4b (two-tailed Student's *t*-test; for *P* values, see the 'Statistics and reproducibility' section in the Methods. The horizontal line denotes mean values. NS, not significant; ****P* < 0.001. **d**, Western blot analysis for UNG1, UNG2, AID and β-actin in CH12 cells with the indicated genotype transduced or not with a retrovirus expressing FAM72A or FAM72A(W125R) and stimulated with TGFβ, IL-4 and anti-CD40 antibody (CIT). **e**, Western blot analysis for UNG1, UNG2 and β-actin in *Fam72a*^{+/+} and *Fam72a*^{-/-} stimulated splenic B cells. **f**, FACS analysis of Ig expression in wild-type CH12 cells and primary B cells transduced with an empty retrovirus (pMX) or expressing UNG2. Representative plots of three experiments are shown (one-tailed Student's *t*-test). The percentage of IgA-expressing cells is indicated and quantified on the right. The horizontal line denotes mean values. For gel source data, see Supplementary Fig. 1.

sustain CSR¹². To examine the contribution of these isoforms relative to FAM72A, we generated *Ung1*^{-/-}, *Ung2*^{-/-}, *Ung1/2*^{-/-}, *Fam72a*^{-/-}*Ung1*^{-/-} and *Fam72a*^{-/-}*Ung2*^{-/-} CH12 cell clones (Extended Data Fig. 4d). While *Ung1*^{-/-} and *Ung2*^{-/-} cells underwent CSR at wild-type levels (Fig. 3b, Extended Data Fig. 4e), CSR in *Ung1/2*^{-/-} CH12 cells was completely

abolished (Fig. 3b, Extended Data Fig. 4e). While *Fam72a*^{-/-}*Ung2*^{-/-} rescued CSR, *Fam72a*^{-/-}*Ung1*^{-/-} did not (Fig. 3b, Extended Data Fig. 4e). This prompted us to analyse the protein levels of UNG1 and UNG2. We found that deficiency in *Fam72a* resulted in a significant and specific upregulation of the UNG2 isoform (Fig. 3d, Extended Data Fig. 4d), which accumulated on chromatin (Extended Data Fig. 4f). This phenotype was also observed in *Fam72a*^{-/-} primary B cells (Fig. 3e) and was not due to changes at the mRNA level (Extended Data Fig. 4g). The upregulation of UNG2 was almost abolished by re-expressing FAM72A but not FAM72A(W125R) (Fig. 3d). In addition, we showed that UNG2 is degraded via a proteasome-dependent mechanism (Extended Data Fig. 4h, i). We concluded that FAM72A specifically regulates the protein level of the UNG2 isoform by controlling its degradation, although the exact mechanism remains to be determined. We therefore assume that the defect in CSR observed in *Fam72a*^{-/-} and *Fam72a*^{-/-}*Ung1*^{-/-} cells is due to the upregulation of UNG2. Consistent with this, overexpression of murine UNG2 (mUNG2) in wild-type CH12 cells and primary splenic B cells (Extended Data Fig. 5a) suppressed CSR (Fig. 3f). This is in contrast to previous reports^{13–15} and we speculate that this discrepancy might be related to differences in the expression levels of UNG2 and/or to the capacity of human UNG2 to reconstitute murine cells. Nevertheless, expression of the HIV-1 VPR protein, which specifically reduces the protein levels of UNG2 (ref. 16), rescued CSR (Extended Data Fig. 5b, c). Moreover, while expression of mUNG2 fused to an auxin-inducible degron¹⁷ (mUNG2^{degron}) only partially rescued CSR in *Ung1/2*^{-/-} CH12 cells (Extended Data Fig. 5d), the addition of auxin completely restored CSR (Extended Data Fig. 5e). Together, it appears that the maintenance of the physiological levels of UNG2 is of critical importance for the efficiency of CSR. Our results suggest that FAM72A, by controlling the level of UNG2, influences the usage of error-prone versus error-free DNA repair in response to AID-induced uracils in DNA.

To determine whether deficiency in *Fam72a* has an influence in the frequency of AID-induced mutations, we sequenced the Sμ switch region in activated CH12 cells. The mutation frequency was reduced by 50% in *Fam72a*^{-/-} CH12 cells (Fig. 4a). Consistent with this, overexpression of mUNG2 in CH12 cells also resulted in a 50% decrease in mutation frequency (Fig. 4a). Overexpression of FAM72A, which had no effect on CSR (Fig. 3a, c), did not lead to a statistically significant difference in mutation frequency (Fig. 4a). To further demonstrate that upregulation of UNG2 has a significant effect on SHM, we made use of an AID-deficient chicken DT40 B cell line, in which all of the pseudo V genes at the light chain locus have been deleted¹⁸ and B cell receptor diversification happens exclusively via SHM when AID is re-expressed. As these cells are IgM⁺, SHM can be assessed by scoring the appearance of IgM⁻ cells after AID expression. The percentage of IgM⁻ cells induced by AID expression was reduced by 2.5-fold when UNG2 was co-expressed with AID (Fig. 4b). To further confirm that deficiency in *Fam72a* results in impaired SHM, we analysed the J_H4 intron in germinal centre B cells isolated from the Peyer's patches of unimmunized mice (Extended Data Fig. 6a). The mutation frequency at the J_H4 intron sequence was drastically reduced in *Fam72a*^{-/-} B cells (Fig. 4c, Extended Data Fig. 6b, c). Consistent with upregulation of UNG2, we found that the percentage of transition mutations at C:G base pairs, which are generated by replication over U:G mismatches, were reduced from 43% in *Fam72a*^{+/+} to 16% in *Fam72a*^{-/-} B cells (Fig. 4c). Conversely, even though the mutation frequency was greatly reduced in *Fam72a*^{-/-} cells, the relative percentage of transversion mutations at C:G base pairs, which are generated through replication over abasic sites, increased from 57% in *Fam72a*^{+/+} to 84% in *Fam72a*^{-/-} B cells (Fig. 4c). This phenotype is diametrically opposed to *Ung1/2*^{-/-} B cells⁶, in which transition mutations at C:G base pairs reaches 95% and transversion mutations at C:G base pairs are suppressed to 5%⁶.

To determine whether the effect of *Fam72a* deficiency on SHM is specific to the *Ig* loci, we performed high-throughput sequencing¹⁹ on three prominent AID off-targets. We found that the mutation frequency

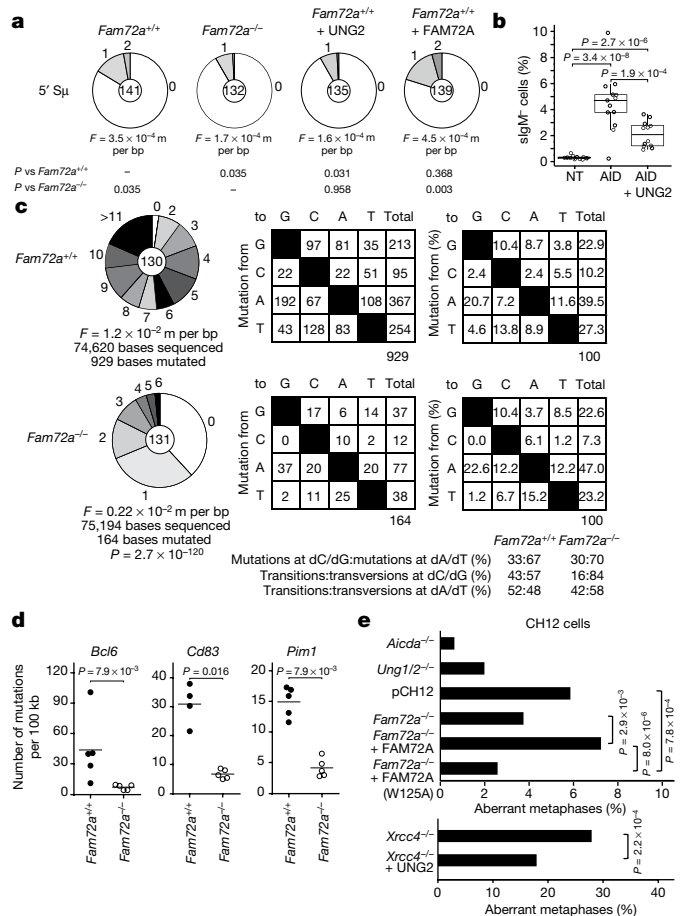


Fig. 4 | Deficiency in *Fam72a* shifts the balance towards error-free DNA repair. **a**, Sμ mutation analysis in stimulated pCH12, *Fam72a*^{-/-} and pCH12 cells transduced with a retrovirus expressing UNG2 or FAM72a. Pie charts depict the proportion of sequences with the indicated number of mutations. The mutation frequency (*F*) is shown below as mutations per base pairs sequenced (m per bp). The total number of sequences analysed is indicated in the centre (two-tailed Student's *t*-test). **b**, IgM expression in slgM⁺ φV *Aicda/ICDA*^{-/-} DT40 cells transduced with a retrovirus expressing AID and/or UNG2. Cells were gated on EGFP and/or mCherry. Twelve independent clones were analysed after 31 days of clonal expansion. For not transduced (NT), minimum = 0.15, maximum = 0.65, median = 0.285, quartile 1 = 0.225, quartile 3 = 0.35, upper whisker bound (UWB) = 0.54125 and lower whisker bound (LWB) = 0.03125. For AID, minimum = 2.44, maximum = 9.9, median = 4.79, quartile 1 = 3.805, quartile 3 = 5.31, UWB = 7.5675 and LWB = 1.5475. For AID + UNG2, minimum = 0.92, maximum = 3.64, median = 2.075, quartile 1 = 1.215, quartile 3 = 2.79, UWB = 5.1525 and LWB = 0 (two-tailed Student's *t*-test). **c**, Analysis of the J_H4 intron mutation in germinal centre B cells from Peyer's patches depicted as in **a**. Mutation profiles are shown on the right. *P* values as determined by SHMtool (χ^2). Data are from five mice of each genotype. **d**, SHM analysis of AID off-targets (Mann–Whitney two-tailed test). **e**, The percentage of aberrant metaphases in stimulated CH12 cells of the indicated genotype harbouring *Igh* locus aberrations (*n* = 470 metaphases (*Aicda*^{-/-}), *n* = 452 (*Ung1/2*^{-/-}), *n* = 887 (pCH12), *n* = 720 (*Fam72a*^{-/-}), *n* = 842 (*Fam72a*^{-/-} + FAM72A) and *n* = 853 (*Fam72a*^{-/-} + FAM72A(W125A))). The bottom panel shows the percentage of aberrant metaphases in *Xrcc4*^{-/-} CH12 cells transduced or not with a retrovirus expressing UNG2 (*n* = 498 metaphases (*Xrcc4*^{-/-}) and *n* = 502 (*Xrcc4*^{-/-} + UNG2)); two independent cell clones were used, see also Supplementary Table 2; two-sided Fisher's exact test.

at these genes was reduced by 3–5-fold in *Fam72a*^{-/-} B cells (Fig. 4d), with a clear inversion in the ratio between transition and transversion mutations at C:G base pairs (Extended Data Fig. 7a). The effect of *Fam72a* deficiency on SHM is therefore not *Ig* loci-specific, suggesting

that FAM72A-mediated regulation of UNG2 could have genome-wide repercussions.

These results support the hypothesis that in FAM72A-deficient B cells, where UNG2 is upregulated, AID-induced uracils are more efficiently excised from DNA by UNG2 and that error-free repair is enforced. Furthermore, enhanced excision of uracil would result in fewer U:G mismatches, rendering mismatch repair less effective in generating mutations at A:T base pairs surrounding the deaminated cytosines. Therefore, the level of UNG2 needs to be tightly controlled to trigger error-prone DNA repair, which is essential for antibody diversification.

The defect in CSR observed in *Fam72a*^{-/-} B cells could be explained by enforced error-free repair due to the upregulation of UNG2, which in turn would lead to the generation of fewer DSBs. In wild-type B cells, a small fraction of AID-induced DSBs can be detected in the form of chromosome breaks or translocations²⁰. To determine whether deficiency in *Fam72a* influences the level of DSBs, we quantified aberrant metaphases harbouring *Igh* locus-specific chromosomal breaks (Extended Data Fig. 7b) in activated *Aicda*^{-/-}, *Ung1/2*^{-/-}, *Fam72a*^{+/+} and *Fam72a*^{-/-} CH12 cells reconstituted (or not) with FAM72A or FAM72A(W125A). We found lower levels of aberrant metaphases in *Fam72a*^{-/-} cells or *Fam72a*^{-/-} cells reconstituted with FAM72A(W125A) than in wild-type and *Fam72a*^{-/-} cells reconstituted with FAM72A (Fig. 4e, Extended Data Fig. 7b, Supplementary Table 2), suggesting that FAM72A-mediated UNG2 upregulation controls the level of AID-induced DSBs during CSR. Consistent with this, overexpression of UNG2 in *Xrcc4*^{-/-} CH12 cells, which accumulate unrepaired DSBs²¹, resulted in a 40% decrease in the frequency of *Igh* DSBs (Fig. 4e). We concluded that the defect in CSR observed in the absence of *Fam72a* is probably due to the inefficient generation of AID-induced DSBs rather than to a perturbation of downstream DSB repair pathways.

Our results are consistent with a model (Extended Data Fig. 8) in which FAM72A interacts with UNG2 to control its physiological level by triggering its degradation. Consequently, deficiency in *Fam72a* leads to the specific upregulation of UNG2 and its accumulation on chromatin. This would enforce uracil excision, resulting in a reduction in the efficiency of SHM and CSR. It is possible that supraphysiological levels of UNG2 in the absence of FAM72A tilt the balance towards error-free base-excision repair by favouring the recruitment of polymerase β^{22–25}. Although the question of how the choice between error-prone and error-free DNA repair is specifically established at *Ig* loci remains to be fully elucidated, we demonstrate here that FAM72A is an UNG2 regulator (and not an AID regulator) and an important player in antibody diversification.

A global and genome-wide shift in the balance from error-free to error-prone DNA repair in response to uracils in DNA has important implications in the onset of cancer. As *Fam72a* is overexpressed in several types of tumours and transformed cell lines⁵ and as *Ung1/2*^{-/-} mice develop B cell lymphomas^{26–28}, it is possible that overexpression of *Fam72a* could suppress the levels of UNG2 leading to inefficient uracil excision, enforcing error-prone DNA repair and making these cells more susceptible to accumulate mutations and hence more prone to tumorigenesis.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-021-04093-y>.

1. Feng, Y., Seija, N., Di Noia, J. M. & Martin, A. AID in antibody diversification: there and back again. *Trends Immunol.* **41**, 586–600 (2020).
2. Methot, S. P. & Di Noia, J. M. Molecular mechanisms of somatic hypermutation and class switch recombination. *Adv. Immunol.* **133**, 37–87 (2017).

3. Stratigopoulou, M., van Dam, T. P. & Guikema, J. E. J. Base excision repair in the immune system: small DNA lesions with big consequences. *Front. Immunol.* **11**, 1084 (2020).
4. Krokan, H. E. et al. Error-free versus mutagenic processing of genomic uracil—relevance to cancer. *DNA Repair* **19**, 38–47 (2014).
5. Guo, C. et al. Ugene, a newly identified protein that is commonly overexpressed in cancer and binds uracil DNA glycosylase. *Cancer Res.* **68**, 6118–6126 (2008).
6. Rada, C. et al. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr. Biol.* **12**, 1748–1755 (2002).
7. Zanotti, K. J. & Gearhart, P. J. Antibody diversification caused by disrupted mismatch repair and promiscuous DNA polymerases. *DNA Repair* **38**, 110–116 (2016).
8. Nakamura, M. et al. High frequency class switching of an IgM⁺ B lymphoma clone CH12F3 to IgA⁺ cells. *Int. Immunol.* **8**, 193–201 (1996).
9. Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR–Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
10. Joung, J. et al. Genome-scale CRISPR–Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**, 828–863 (2017).
11. KOMP strain detail sheet. *MMRRC* <https://www.mmrrc.org/catalog/sds.php?mmrrc_id=47653>
12. Sarno, A. et al. Uracil-DNA glycosylase UNG1 isoform variant supports class switch recombination and repairs nuclear genomic uracil. *Nucleic Acids Res.* **47**, 4569–4585 (2019).
13. Begum, N. A. et al. Uracil DNA glycosylase activity is dispensable for immunoglobulin class switch. *Science* **305**, 1160–1163 (2004).
14. Di Noia, J. M. et al. Dependence of antibody gene diversification on uracil excision. *J. Exp. Med.* **204**, 3209–3219 (2007).
15. Yousif, A. S., Stanlie, A., Mondal, S., Honjo, T. & Begum, N. A. Differential regulation of S-region hypermutation and class-switch recombination by noncanonical functions of uracil DNA glycosylase. *Proc. Natl Acad. Sci. USA* **111**, E1016–E1024 (2014).
16. Eldin, P. et al. Impact of HIV-1 Vpr manipulation of the DNA repair enzyme UNG2 on B lymphocyte class switch recombination. *J. Transl. Med.* **18**, 310 (2020).
17. Holland, A. J., Fachinetti, D., Han, J. S. & Cleveland, D. W. Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. *Proc. Natl Acad. Sci. USA* **109**, E3350–E3357 (2012).
18. Arakawa, H., Saribasak, H. & Buerstedde, J. M. Activation-induced cytidine deaminase initiates immunoglobulin gene conversion and hypermutation by a common intermediate. *PLoS Biol.* **2**, E179 (2004).
19. Martin, O. A. et al. Detecting rare AID-induced mutations in B-lineage oncogenes from high-throughput sequencing data using the detection of minor variants by error correction method. *J. Immunol.* **201**, 950–956 (2018).
20. Dev, H. et al. Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells. *Nat. Cell Biol.* **20**, 954–965 (2018).
21. Yan, C. T. et al. IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* **449**, 478–482 (2007).
22. Krokan, H. E. & Bjoras, M. Base excision repair. *Cold Spring Harb. Perspect. Biol.* **5**, a012583 (2013).
23. Krokan, H. E., Drablos, F. & Slupphaug, G. Uracil in DNA—occurrence, consequences and repair. *Oncogene* **21**, 8935–8948 (2002).
24. Wu, X. & Stavnezer, J. DNA polymerase beta is able to repair breaks in switch regions and plays an inhibitory role during immunoglobulin class switch recombination. *J. Exp. Med.* **204**, 1677–1689 (2007).
25. Schrader, C. E., Guikema, J. E., Wu, X. & Stavnezer, J. The roles of APE1, APE2, DNA polymerase β and mismatch repair in creating S region DNA breaks during antibody class switch. *Phil. Trans. R. Soc. B* **364**, 645–652 (2009).
26. Nilsen, H. et al. Gene-targeted mice lacking the Ung uracil-DNA glycosylase develop B-cell lymphomas. *Oncogene* **22**, 5381–5386 (2003).
27. Andersen, S. et al. Monoclonal B-cell hyperplasia and leukocyte imbalance precede development of B-cell malignancies in uracil-DNA glycosylase deficient mice. *DNA Repair* **4**, 1432–1441 (2005).
28. Nilsen, H., An, Q. & Lindahl, T. Mutation frequencies and AID activation state in B-cell lymphomas from Ung-deficient mice. *Oncogene* **24**, 3063–3066 (2005).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021

Methods

CH12^{Cas9} cell generation

CH12 cells were transduced with retroviral supernatants obtained by transfecting Bosc23 cells with a retrovirus (pMX-26) expressing a Cas9-mCherry-P2A-hygromycin cassette (Extended Data Fig. 1a). After 72 h of hygromycin selection (300 µg/ml; Sigma), mCherry⁺ cells were sorted into 96-well plates using a cell-sorter FACS ARIA Fusion (Becton Dickinson) and cultured for 10 days with hygromycin (300 µg/ml; Sigma). Individual clones were then tested for their ability to express Cas9 by western blot (Extended Data Fig. 1a) and their ability to undergo efficient CSR after stimulation (Extended Data Fig. 1a). To test the functionality of Cas9, clones were transfected with gRNAs known to be effective (see Supplementary Table 3 for gRNA sequences). Three days later, PCR was performed to identify clones with Cas9-mediated deletions (Extended Data Fig. 1a). Clone #11, which had the highest Cas9 activity and displayed a CSR efficiency similar to the parental CH12 cell line, was selected to conduct the screen.

Subcloning of the mBrie gRNA library

The mouse Brie CRISPR knockout pooled library (lentiCRISPRv2 backbone; a gift from D. Root and J. Doench (Addgene 73633)) was subcloned into the pMX-28 retroviral vector (Extended Data Fig. 1b) as previously described¹⁰ to generate the pMX-mBrie gRNA library. In brief, the U6p-gRNA-scaffold cassette was amplified by PCR (see Supplementary Table 3 for primers), digested with BamHI and NotI restriction enzymes and ligated into pMX-28 (Extended Data Fig. 1b). Stbl4-electrocompetent cells (Thermo Fisher) were electroporated using Eporator Eppendorf (Program 1-1700V), plated on ampicillin agar plates and incubated at 37 °C for 15 h. Plasmid DNA was extracted and the U6-gRNA-scaffold cassette was amplified by PCR to add sequencing adaptors and barcodes (see Supplementary Table 3 for primers) and then analysed by deep sequencing to determine gRNA representation and library uniformity as previously described¹⁰.

CRISPR-Cas9 knockout screen

CH12^{Cas9} cells were transduced with the pMX-mBrie gRNA library in quadruplicate at a multiplicity of infection (MOI) of 0.3 and with a 300× coverage, as previously described¹⁰. Transduced cells were selected with puromycin (0.5 µg/ml; Sigma) and hygromycin (300 µg/ml; Sigma) for 15 days. Selected cells were then induced to undergo CSR with TGFβ (1 ng/ml; R&D Systems Europe), IL-4 (5 ng/ml; Peprotech) and an anti-CD40 antibody (200 ng/ml; eBioscience) for 72 h. IgM⁺ and IgA⁺ cells were sorted using anti-IgM and anti-IgA coupled magnetic beads (Miltenyi). Population purity was assessed by flow cytometry (Extended Data Fig. 1c) using a Fortessa flow cytometer (Becton Dickinson). Genomic DNA from 24 million cells was extracted using phenol/chloroform and subjected to PCR (one cycle at 95 °C for 2 min; 28–31 cycles at 95 °C for 15 s, 65 °C for 20 s, 72 °C for 30 s; and one cycle at 72 °C for 3 min) to amplify gRNA sequences using staggered primers having Illumina adaptors and barcodes (see Supplementary Table 3 for primers). Multiplexed samples were submitted to high-throughput sequencing (1 × 50 bp) on an Illumina HiSeq4000 sequencer at the GenomEast sequencing platform of Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC). gRNA counts were extracted from raw data using the PoolQ software from the Broad Institute. gRNA counts were normalized according as previously described¹⁰. Ranking of candidate genes was performed using MaGeCK²⁹. Candidate genes ($P < 0.05$) and having more than two effective gRNAs were selected. Gene ontology analysis of genes identified in the screen was performed with Enrichr database and statistical analysis was performed using two-tailed Fisher's exact test.

Cell culture

Bosc23 cells were cultured in DMEM supplemented with glucose (4.5 g/l), 10% of heat-inactivated fetal calf serum (FCS),

penicillin-streptomycin (100 U/ml) and sodium pyruvate (1 mM). CH12 cells and primary B cells were cultured in RPMI supplemented with 10% FCS, HEPES (10 mM), penicillin-streptomycin (100 U/ml), sodium pyruvate (1 mM) and β-mercaptoethanol (50 µM). CH12 cells were cultured in the presence of cycloheximide (CHX; 10 µg/ml; Fisher Scientific) and MG132 (10 µM; R&D Systems) and auxin (IAA; 5–500 µM; Sigma-Aldrich). slgM⁺ φV *Aicda*^{-/-} DT40 cells were cultured in RPMI 1640 supplemented with 9% FCS, 1% chicken serum, β-mercaptoethanol (50 µM) and penicillin-streptomycin (100 U/ml) at 37 °C in 5% CO₂. HEK293T cells were cultured in DMEM supplemented with 10% FCS and penicillin-streptomycin (100 U/ml). All cell lines used were mycoplasma negative.

CSR assays

CH12 cells were cultured for 72 h in the presence of TGFβ (1 ng/ml; R&D Systems Europe), IL-4 (5 ng/ml; Peprotech) and an anti-CD40 antibody (200 ng/ml; eBioscience). Cells were then stained with an anti-IgA-PE antibody (1/500; Southern Biotech) to assess CSR by flow cytometry. Before analysis, DAPI was added to discriminate dead cells. Samples were analysed using a Fortessa flow cytometer (Becton Dickinson) and the FlowJo software.

Primary B cell cultures

Splenic resting B cells were purified using anti-CD43 magnetic beads (Miltenyi), labelled with CFSE (Thermo Fisher Scientific) and cultured for 4 days with a combination of LPS (25 µg/ml; Sigma-Aldrich), IL-4 (25 µg/ml; Peprotech), anti-IgD-dextran (6 ng/ml; Fina Biosolutions), IFNγ (100 ng/ml; Peprotech), IL-5 (5 ng/ml; R&D Systems), TGFβ (3 ng/ml; R&D Systems) and retinoic acid (RA; 0.3 ng/ml; Sigma-Aldrich).

RT-qPCR

RNA and cDNA were prepared using standard techniques. Quantitative PCR (qPCR) was performed in triplicates using Roche LightCycler 480 Probes Master mix Universal Probe Library (UPL) in combination with appropriate UPL probes (Supplementary Table 3) or QuantiTect SYBR Green Master mix (Qiagen). Transcript quantities were calculated relative to standard curves and normalized to *Hprt* or *Igb* (also known as *Cd79b*) mRNA.

Retroviral transductions

CH12 cells were transduced with retroviral supernatants obtained by transfecting Bosc23 cells with an empty retrovirus (pMX-PIE; puromycin-IRES-EGFP) or expressing mFAM72A (pMX-FAM72A), mFAM72A(W125R) (pMX-mFAM72A(W125R)), mFAM72A(W125A) (pMX-FAM72A(W125A)), mUNG2 (pMX-mUNG2), Ugi (pMX-Ugi), HIV1-VPR (pMX-VPR), mUNG2^{Degron} (pMX-mUNG2^{Degron}) or TIR1 (pCH-TIR1). Transduced cells were then selected with puromycin (1 µg/ml) and/or hygromycin (300 µg/ml; Sigma) for 7 days and submitted to CSR assays. slgM⁺ φV *Aicda*^{-/-} DT40 cells were transduced with retroviral supernatants obtained by transfecting HEK293T cells with an empty retrovirus (pMX-PIE; puromycin-IRES-EGFP) or expressing mUNG2 (pMX-mUNG2) and with a retrovirus expressing mAID (pCH-8, mCherry and hygromycin).

SHM analysis on DT40 cells

Transduced slgM⁺ φV *Aicda*^{-/-} DT40 cells were seeded (5 × 10⁵ cells per well) in 24-well plates in 1 ml of medium. After 24 h, cells were transferred to 96-well plates in culture medium supplemented with puromycin (0.25 µg/ml) and hygromycin (300 µg/ml). After 48 h, cells were cultured under limiting dilution conditions to obtain single-cell clones. After 2 weeks, antibiotic-resistant colonies were picked and expanded for 2 additional weeks. The frequency of slgM-loss variants was then determined by flow cytometry after staining with anti-chicken IgM-PE antibody (Southern Biotech).

Co-immunoprecipitation

Whole-cell extracts prepared from wild-type and *Fam72a*^{-/-} CH12 cells transduced with a retrovirus expressing FAM72A-Flag or FAM72A(W125R)-Flag were immunoprecipitated with an anti-Flag antibody coupled to agarose beads (M2; Sigma) and analysed by western blot.

Generation of CH12 knockout clones

CH12 cells were transfected by electroporation using the Neon transfection System (Thermo Fisher) with a plasmid expressing one or two gRNAs targeting a critical exon (see Supplementary Table 3 for gRNA sequences) and co-expressing the high-fidelity Cas9 nuclease (Cas9-HF1)³⁰ coupled to EGFP. Twenty-four hours after transfection, individual EGFP-positive cells were sorted into 96-well plates using a cell-sorter FACS ARIA Fusion (Becton Dickinson) and cultured for 10 days. Clones were then genotyped by PCR and sequencing.

μ and JH4i somatic hypermutation analysis

For μ mutation analysis, genomic DNA was extracted from CH12 cells cultured for 3 days with TGFβ (1 ng/ml; R&D Systems Europe), IL-4 (5 ng/ml; Peprotech) and an anti-CD40 antibody (200 ng/ml; eBioscience). For JH4i mutation analysis, genomic DNA was extracted from B cells isolated from the Peyer's patches and sorted by flow cytometry using anti-B220-PE-Cy7 (eBiosciences), anti-Fas-PE (BD Pharmingen) and anti-GL7-pacific blue (BioLegend) antibodies. μ and JH4i sequences were amplified by PCR (one cycle at 98 °C for 30 s; 35 cycles at 98 °C for 10 s, 70 °C for 10 s and 72 °C for 30 s; and one cycle at 72 °C for 5 min) using the Q5 polymerase (New England BioLabs) and cloned into pUC57 using the MEGAWHOP³¹ method. Inserts were sequenced (Sanger sequencing) using the M13 forward universal primer. Sequences were aligned with Lasergene (DNASTAR) and analysed with the SHMTool³² server.

SHM analysis of AID off-targets

Genomic regions within the *Bcl6*, *Cd83* and *Pim1* genes were amplified by PCR using previously described conditions and primers¹⁹ from genomic DNA obtained from germinal centre B cells isolated from the Peyer's patches of unimmunized mice. Next-generation sequencing was performed according to the user guide Ion Xpress Plus gDNA Fragment Library Preparation (Life Technologies) from PCR products (100 ng each), and libraries were sequenced on an Ion torrent S5 system. Raw sequencing reads obtained in fastq format were analysed using the DeMinEr pipeline¹⁹. Samples from *Aicda*^{-/-} B cells were used to determine the baseline mutation background and filter out mutations due to sequencing errors. Python scripts used in this study are available in a public repository under GNU General Public License v3.0 (<https://github.com/fboyerCRIBL/deminer>) and can also be provided upon demand from the corresponding author.

Cloning

All gRNA/Cas9-HF1 plasmids were generated through golden gate cloning³³. cDNAs were cloned into the pMX-PIE retrovirus through SLICE³⁴.

Mice

Fam72a^{-/-} mice (C57BL6) were generated by the knockout mouse program (KOMP)¹¹. Mice were bred under pathogen-free (SPF) conditions with standard dark/light cycles, ambient temperature and humidity. *Fam72a*^{-/-} mice gave offspring at Mendelian ratios and showed no obvious deleterious phenotype. In all experiments 8–12-week-old age-matched littermates (males and females) were used. Animal work was performed under protocols approved by the IGBMC's ethics committee (ComEth/APAFIS#23104-2019112915476749) and by the French Ministry of Higher Education, Research and Innovation (D6721837).

Igh FISH

Metaphases were prepared using standard procedures³⁵. DNA FISH on metaphase spreads was performed as previously described³⁵ using BAC probes RP24-134G24 (5' *Igh* C) and RP24-386J17 (3' *Igh* V) and X-Cyting mouse chromosome 12 (Orange) paint from MetaSystems. Metaphases were imaged using a ZEISS AxioImager.Z2 microscope and the Metafer automated capture system (MetaSystems), and counted manually.

Western blot analysis

Protein extracts were prepared using standard techniques. Proteins were separated by SDS-PAGE using gradient gels (4–12%; Invitrogen), transferred to PVDF membranes (Immobilon; Millipore) and analysed using anti-AID³⁶, anti-β-actin (Sigma), anti-UNG1/2 (a gift from B. Kavli), anti-Nbs1 (a gift from A. Nussenzweig), anti-histone H3 (Abcam) and anti-Flag (Sigma) antibodies. Cell fractionation experiments were performed as previously described³⁷.

B cell development

Total splenic or bone marrow cells were labelled with suitable antibodies and analysed on an FACS-Fortessa flow cytometer (Becton Dickinson).

Antibodies

Flag (1/1,000; M2, Sigma), IgA-biotin (1/500; Southern Biotech), UNG (a gift from B. Kavli; 2 μg/ml), AID (1/10,000; IGBMC), β-actin (1/200,000; Sigma), IgG1-biotin (1/500; BD Pharmingen), IgG2a-biotin (1/250; BioLegend), IgG2b-biotin (1/500; BioLegend), IgG3-biotin (1/250; BD Pharmingen), IgA-PE (1/500; Southern Biotech), GL7-pacific blue (1/200; BioLegend), B220-PE-Cy7 (1/100; eBiosciences), Fas (CD95)-PE (1/100; BD Pharmingen), streptavidin-PE (1/500; Jackson ImmunoResearch), histone H3 (1/10,000; Abcam), Nbs1 (a gift from A. Nussenzweig; 1/2,000), anti-CD40 (200 ng/ml; eBioscience) and anti-chicken IgM-PE (1/500; Southern Biotech).

Statistics and reproducibility

For Fig. 3c, *P* values are pCH12 versus pCH12 + FAM72A = 0.45, pCH12 versus *Fam72a*^{-/-} = 2.2×10^{-8} , pCH12 versus *Fam72a*^{-/-} + pMX = 1.9×10^{-7} , pCH12 versus *Fam72a*^{-/-} + FAM72A = 0.15, pCH12 versus *Fam72a*^{-/-} + FAM72A(W125R) = 8.85×10^{-8} , pCH12 versus *Fam72a*^{-/-} + FAM72A(W125A) = 4.81×10^{-9} , pCH12 versus *Fam72a*^{-/-} + Ugi = 7.31×10^{-9} , *Fam72a*^{-/-} versus *Fam72a*^{-/-} + pMX = 0.77, *Fam72a*^{-/-} versus *Fam72a*^{-/-} + FAM72A = 1.59×10^{-5} , *Fam72a*^{-/-} versus *Fam72a*^{-/-} + FAM72A(W125R) = 0.73, *Fam72a*^{-/-} versus *Fam72a*^{-/-} + FAM72A(W125A) = 0.79, *Fam72a*^{-/-} versus *Fam72a*^{-/-} + Ugi = 2.43×10^{-6} .

For Extended Data Fig. 5c, *P* values are pCH12 versus *Fam72a*^{-/-} = 3×10^{-4} ; pCH12 versus *Fam72a*^{-/-} + pMX-VPR = 4.3×10^{-1} ; pCH12 versus *Fam72a*^{-/-} + FAM72A = 5.6×10^{-1} ; pCH12 versus *Fam72a*^{-/-} + FAM72A + pMX-VPR = 4.9×10^{-2} ; pCH12 versus *Fam72a*^{-/-} + FAM72A(W125A) = 1.8×10^{-3} ; pCH12 versus *Fam72a*^{-/-} + FAM72A(W125A) + pMX-VPR = 6.4×10^{-1} ; *Fam72a*^{-/-} versus *Fam72a*^{-/-} + pMX-VPR = 1.1×10^{-2} ; *Fam72a*^{-/-} + FAM72A versus *Fam72a*^{-/-} + FAM72A + pMX-VPR = 2×10^{-1} ; *Fam72a*^{-/-} + FAM72A(W125A) versus *Fam72a*^{-/-} + FAM72A(W125A) + pMX-VPR = 4.3×10^{-3} .

For Extended Data Fig. 5d, *P* values are pCH12 – IAA versus pCH12 + 500 μM IAA = 3.8×10^{-1} ; pCH12 – IAA versus *Fam72a*^{-/-} – IAA = 7.7×10^{-6} ; pCH12 – IAA versus *Fam72a*^{-/-} + 500 μM IAA = 7.8×10^{-6} ; pCH12 – IAA versus *Ung1/2*^{-/-} – IAA = 1.9×10^{-6} ; pCH12 – IAA versus *Ung1/2*^{-/-} + 500 μM IAA = 1.9×10^{-6} ; pCH12 – IAA versus *Ung1/2*^{-/-} + UNG2^{Degron} + TIR1 – IAA = 3.9×10^{-4} ; pCH12 – IAA versus *Ung1/2*^{-/-} + UNG2^{Degron} + TIR1 + 5 μM IAA = 3.2×10^{-2} ; pCH12 – IAA versus *Ung1/2*^{-/-} + UNG2^{Degron} + TIR1 + 500 μM IAA = 6.8×10^{-1} ; *Fam72a*^{-/-} – IAA versus *Fam72a*^{-/-} + 500 μM IAA = 3.9×10^{-2} ; *Ung1/2*^{-/-} – IAA versus *Ung1/2*^{-/-} + 500 μM IAA = 8.9×10^{-5} ; *Ung1/2*^{-/-} + UNG2^{Degron} + TIR1 – IAA versus *Ung1/2*^{-/-} + UNG2^{Degron} + TIR1 + 5 μM IAA = 1.4×10^{-5} ; *Ung1/2*^{-/-} + UNG2^{Degron} + TIR1 – IAA versus *Ung1/2*^{-/-} + UNG2^{Degron}

Article

+TIRI + 500 μ M IAA = 6.8×10^{-5} ; *Ung1/2^{-/-}* + UNG2^{Degron} + TIRI + 5 μ M IAA versus *Ung1/2^{-/-}* + UNG2^{Degron} + TIRI + 500 μ M IAA = 4.3×10^{-3} .

For Fig. 3d, e, data presented are representative of three and two independent experiments, respectively.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The CRISPR–Cas9 screen high-throughput sequencing data are available at the Gene Expression Omnibus (GSE184145). Source data for Figs. 1–4 can be found in Supplementary Information. Source data are provided with this paper.

29. Li, W. et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).
30. Kleinstiver, B. P. et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
31. Miyazaki, K. MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. *Methods Enzymol.* **498**, 399–406 (2011).
32. Maccarthy, T., Roa, S., Scharff, M. D. & Bergman, A. SHMTool: a webserver for comparative analysis of somatic hypermutation datasets. *DNA Repair* **8**, 137–141 (2009).
33. Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS ONE* **4**, e5553 (2009).
34. Motohashi, K. A simple and efficient seamless DNA cloning method using SLiCE from *Escherichia coli* laboratory strains and its application to SLiP site-directed mutagenesis. *BMC Biotechnol.* **15**, 47 (2015).
35. Lescale, C. et al. RAG2 and XLF/Cernunnos interplay reveals a novel role for the RAG complex in DNA repair. *Nat. Commun.* **7**, 10529 (2016).
36. Jeevan-Raj, B. P. et al. Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *J. Exp. Med.* **208**, 1649–1660 (2011).
37. Holden, P. & Horton, W. A. Crude subcellular fractionation of cultured mammalian cell lines. *BMC Res. Notes* **2**, 243 (2009).

Acknowledgements We thank A. Frey for help with gRNA library subcloning and the CRISPR–Cas9 screen; B. Kavli for the anti-UNG1/2 antibody; L. Brino for help with gRNA library

subcloning and CRISPR–Cas9 screen analysis; B. Jost, C. Thibault-Carpentier and D. Plassard for gRNA high-throughput sequencing on the GenomEast platform of IGBMC; D. Dembele, S. Sarnataro and N. Molina for help with Python scripts and gRNA extraction and counting; B. Heller for cell culture; C. Ebel and M. Philipps for cell sorting in the FlowCytometry Facility of IGBMC; M. Li for the GL-7 antibody, M. Selloum, D. Ali-Hadji and I. Gonçalves Da Cruz for animal care; W. Yu for generating *Xrcc4^{-/-}* CH12 cells; and C. Goujon for the HIV1-VPR cDNA. J.M. and M.T. were supported by the Ministère de l'Enseignement Supérieur et de la Recherche, France, and the Fondation ARC. M.R. was supported by La Ligue Contre le Cancer. This study was supported by grants from the Fondation Recherche Médicale (Equipe FRM EQU201903007818 to B.R.-S.-M.), Institut National du Cancer (INCa_13852 to L.D. and B.R.-S.-M.), Fondation ARC (ARCPJA32020060002061 to B.R.-S.-M.), the Ligue Nationale contre le Cancer (Equipe labellisée to L.D.) and by the grant ANR-10-LABX-0030-INRT, a French state fund managed by the Agence Nationale de la Recherche under the program Investissements d'Avenir labelled ANR-10-IDEX-0002-02.

Author contributions B.R.-S.-M. conceived and designed the study. M.R. generated CH12^{Cas9} cells, subcloned the gRNA library and performed the CRISPR–Cas9 screen with I.R. Screen data were analysed by B.R.-S.-M. and M.R. All CH12 knockout cell lines were generated and characterized by M.R. and J.M. with the help of A.A. The majority of experiments in CH12 cells and in *Fam72a^{-/-}* mice were performed by M.R. and J.M. *Igh* FISH experiments were performed and analysed by C.L. All constructs were cloned by M.R. and J.M., with help from V.H. V.H. performed the nuclear/chromatin fractionations and provided assistance for western blot analysis. M.R. and J.M. performed the SHM experiments in CH12 cells with the help of V.H. M.R. and J.M. performed SHM analysis in germinal centre B cells, with the help of A.A. O.M. generated the AID off-target high-throughput sequencing libraries, which were sequenced and analysed by M.T. and E.P. K.C. performed and analysed SHM in DT40 cells. A.-S.T.-C. performed mRNA sequencing and analysed the data. B.L. generated VPR lentiviral particles. F.J. performed immunohistochemistry on spleen sections. E.P., K.T., M.C., S.G.C. and E.S. oversaw experiments. M.R., J.M., E.S. and B.R.-S.-M. wrote the manuscript. M.R., J.M., C.L., E.S., L.D. and B.R.-S.-M. edited the manuscript. The overall research was directed by L.D. and B.R.-S.-M.

Competing interests The authors declare no competing interests.

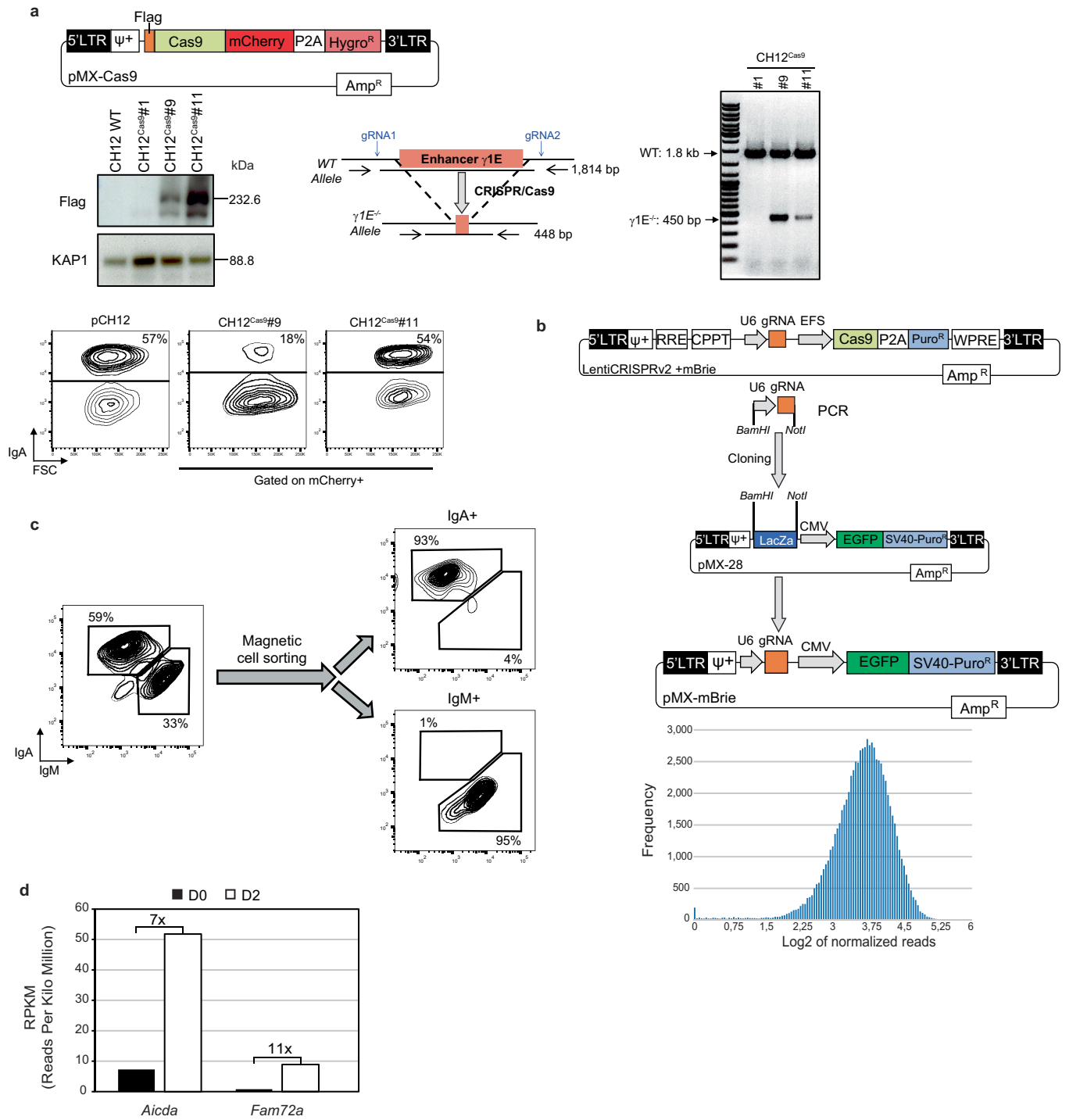
Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41586-021-04093-y>.

Correspondence and requests for materials should be addressed to Bernardo Reina-San-Martin.

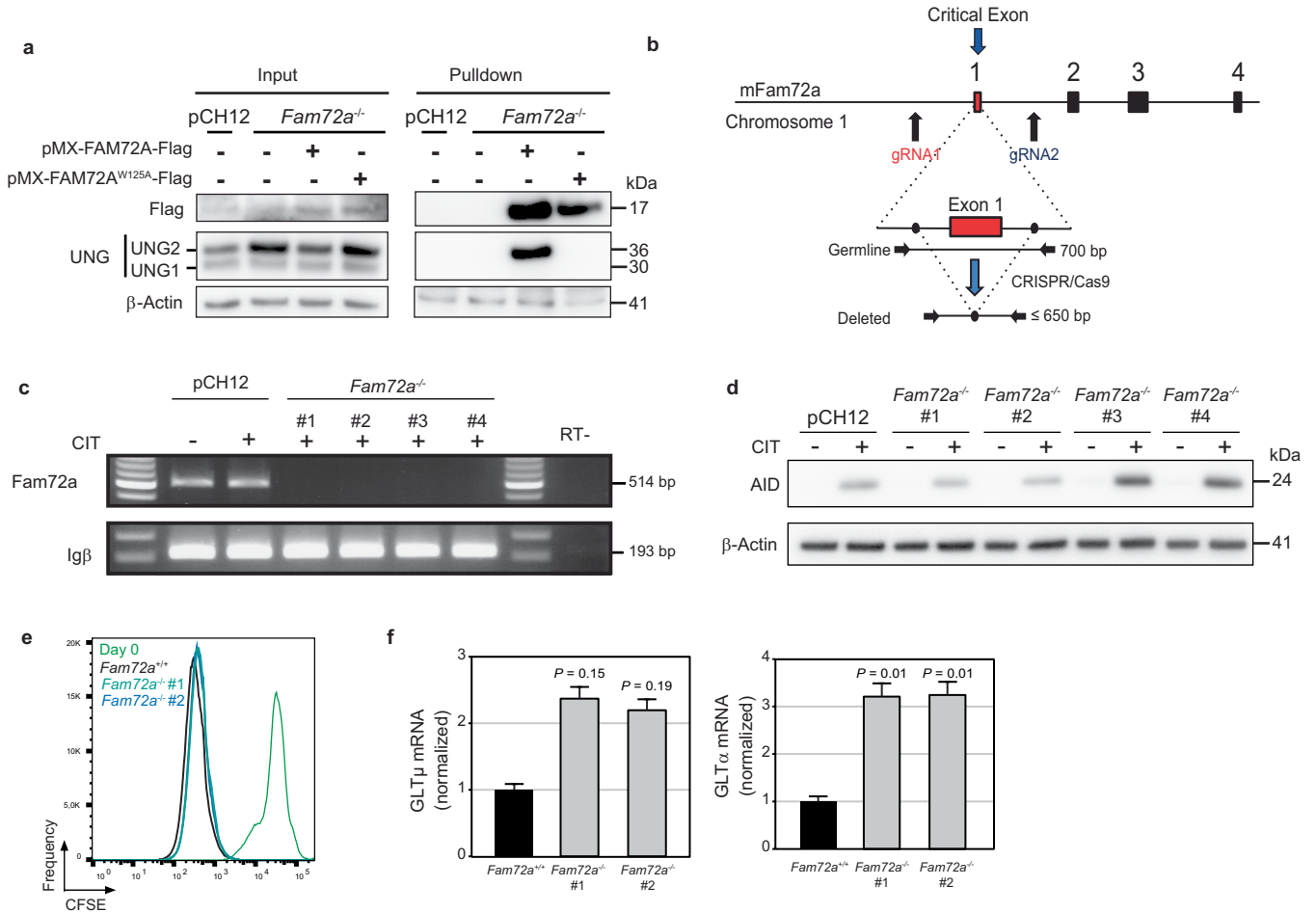
Peer review information Nature thanks Uttiya Basu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>.



Extended Data Fig. 1 | Setting up the CRISPR/Cas9 genome-wide knock-out screen. a, Generation and validation of Cas9-expressing (CH12^{Cas9}) cells. Clones were verified by Western blot for Cas9 expression, by PCR for Cas9-induced deletion of the enhancer $\gamma 1$ ($\gamma 1E$) and by flow cytometry for CSR. Clone #11 was chosen for the screen. Blots and data of CSR assay are representative of 3 experiments. **b**, The U6-gRNA sequence was amplified by PCR from the mBrie library and subcloned into pMX-28 using BamHI and NotI

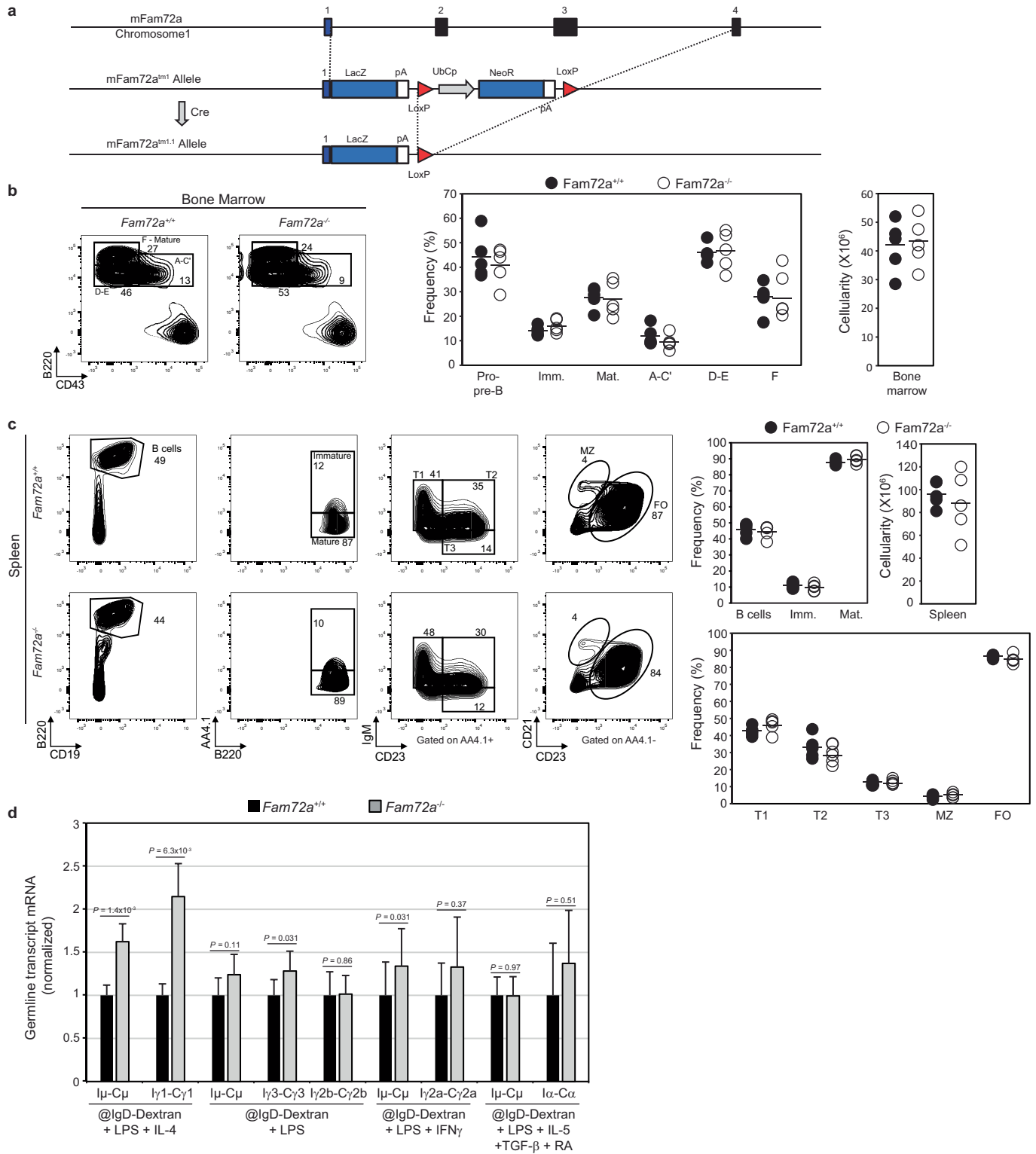
restriction enzymes to generate the pMX-mBrie gRNA library. gRNA representation was analyzed by high-throughput sequencing. **c**, Purity of IgM⁺ and IgA⁺ sorted populations was verified by flow cytometry. **d**, Number of reads at the *Aicda* and *Fam72a* genes in wild-type splenic B cells before and after 48h in culture with LPS and IL-4, as determined by mRNA-Seq. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 2 | FAM72A interacts with UNG2 and *Fam72a*^{-/-} CH12 cells have no defects in proliferation, AID expression or switch region transcription.

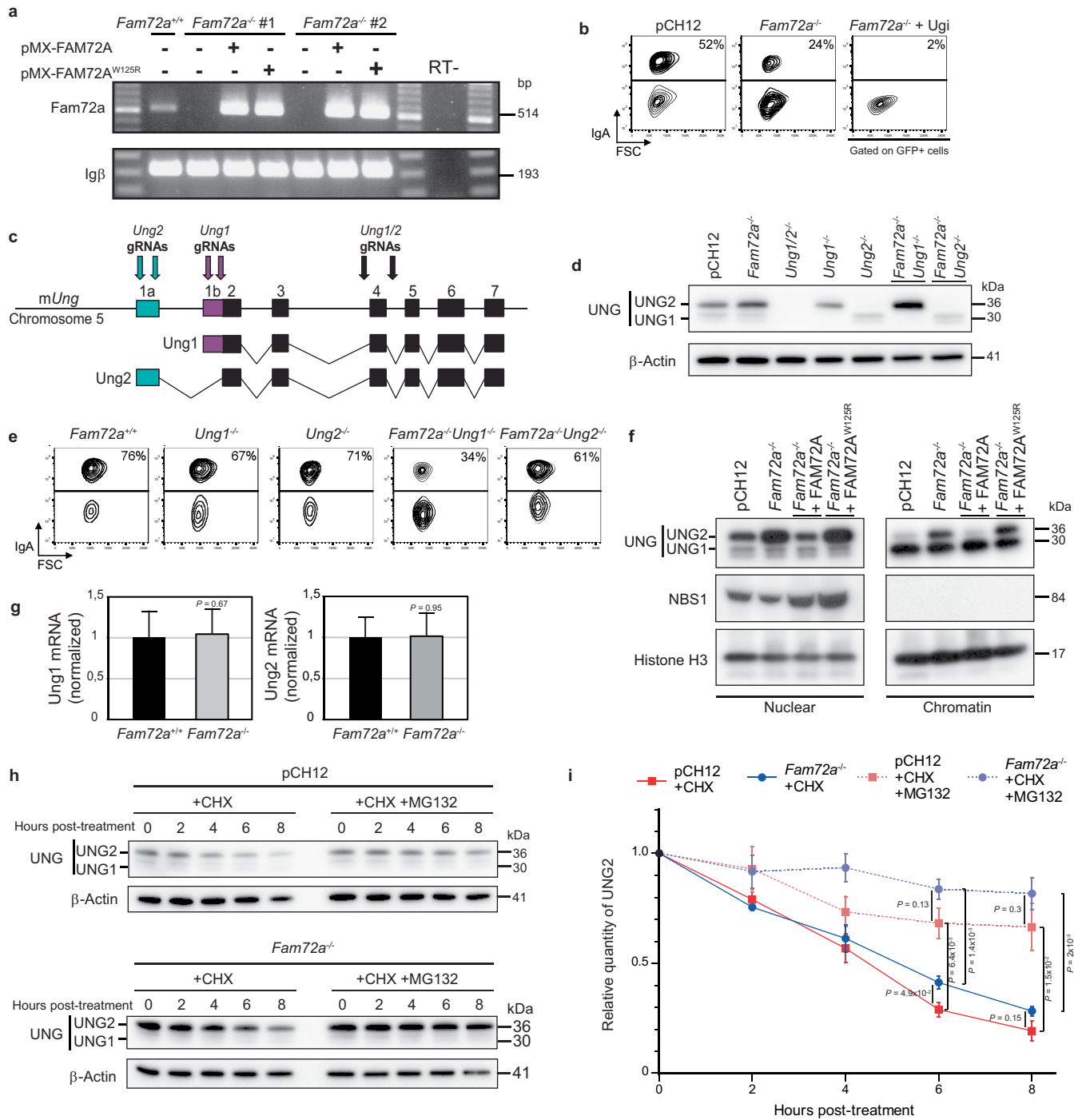
a, Western blot analysis for Flag, UNG (UNG1 and UNG2) and β-Actin in wildtype and *Fam72a*^{-/-} CH12 cells transduced with a retrovirus expressing FAM72A-Flag or FAM72A^{W125A}-Flag before (Input) or after immunoprecipitation with a Flag-specific antibody (Pulldown). **b**, Schematic representation of the murine *Fam72a* locus and location of the gRNAs used to generate *Fam72a*^{-/-} clones using CRISPR/Cas9-HF1. **c**, RT-PCR analysis for *Fam72a* and *Igβ* in *Fam72a*^{+/+} and *Fam72a*^{-/-} CH12 cells cultured or not for 3 days with TGFβ, anti-CD40 antibody and IL-4 (CIT). Data are representative of three experiments. **d**, Western blot analysis for AID and β-actin in wild-type (pCH12

and *Fam72a*^{-/-} CH12 cells cultured or not for 3 days with TGFβ, anti-CD40 antibody and IL-4 (CIT). **e**, CFSE dye-dilution analysis by flow cytometry in *Fam72a*^{+/+} and two independent *Fam72a*^{-/-} CH12 cell clones cultured for 3 days with TGFβ, IL-4 and anti-CD40 antibody. Data are representative of three experiments. **f**, RT-qPCR analysis for GLTμ and GLTα in *Fam72a*^{+/+} and *Fam72a*^{-/-} CH12 cells cultured for 3 days with TGFβ, IL-4 and anti-CD40 antibody. Triplicates were normalized to the abundance of *Igβ* and are expressed relative to *Fam72a*^{+/+}, set as 1. Statistical significance was determined by a two-tailed Student's *t*-test. Data are presented as mean ± s.d. and are representative of 3 experiments. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 3 | Robust B cell development and switch region transcription in *Fam72a*^{-/-} mice. **a**, Schematic representation of the murine *Fam72a* locus and strategy for the generation of *Fam72a*^{-/-} mouse model. **b, c**, Flow cytometry and cellular analysis of B cell populations in the bone marrow (**b**) or the spleen (**c**) from *Fam72a*^{+/+} and *Fam72a*^{-/-} mice, using the indicated cell surface markers. The data are representative of 5 mice per genotype. Horizontal line: mean values. **d**, Real-time qPCR analysis for germline

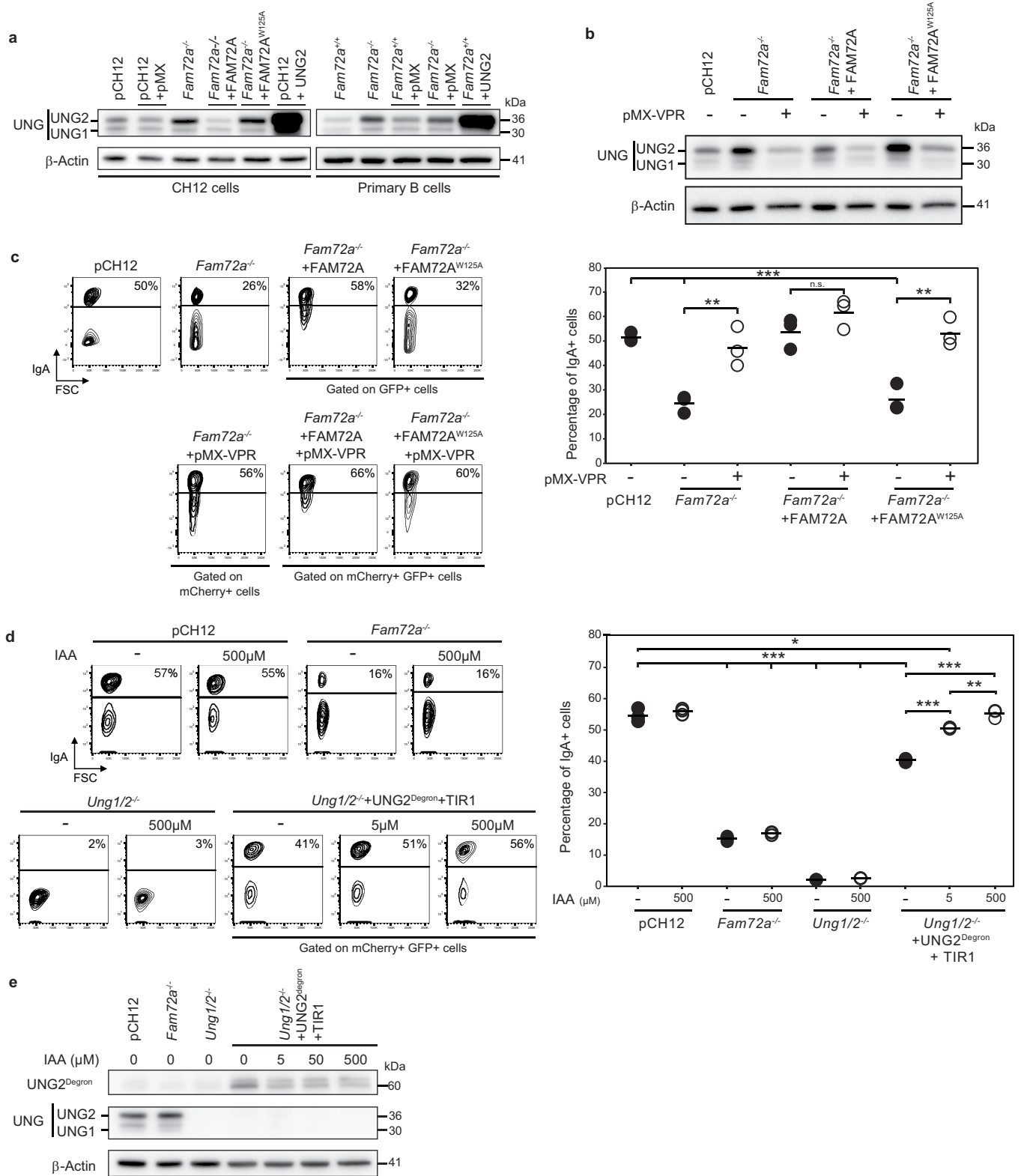
transcripts at donor (GLTμ) and acceptor switch regions (GLTγ3, GLTγ1, GLTγ2b, GLTγ2a and GLTα) in *Fam72a*^{+/+} and *Fam72a*^{-/-} splenic B cells cultured for 96h as in **c**. Expression is normalized to Igβ and is presented relative to expression in *Fam72a*^{+/+} B cells, set as 1. Mean of three mice per genotype + SEM were calculated following the rules for error propagation while calculating a ratio. Statistical analysis was performed using two-tailed Student's *t*-test.



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | FAM72A specifically controls UNG2 protein levels and its accession to chromatin via a proteasome-dependent mechanism. **a**, RT-PCR analysis for *Fam72a* and *Igβ* in *Fam72a*^{+/+} and *Fam72a*^{-/-} CH12 cells transduced (or not) with a retrovirus expressing FAM72A, FAM72A^{W125R} or FAM72A^{W125A}. Data are representative of three experiments. **b**, Flow cytometry analysis of IgA expression in *Fam72a*^{+/+} and *Fam72a*^{-/-} CH12 cells expressing (or not) an UNG inhibitor (Ugi) and cultured for 3 days with TGFβ, anti-CD40 antibody and IL-4. Representative plots are shown. The percentage of IgA-expressing cells is indicated. **c**, Schematic representation of the murine *Ung* locus and location of the gRNAs targeting *Ung2* (exon 1a; blue), *Ung1* (exon 1b; purple) or *Ung1/2* (exon 4; black) used to generate *Ung1*^{-/-}, *Ung2*^{-/-}, *Ung1*^{-/-} *Ung2*^{-/-}, *Fam72a*^{-/-} *Ung1*^{-/-} and *Fam72a*^{-/-} *Ung2*^{-/-} CH12 cell clones using CRISPR/Cas9-HF1. Note, the gRNAs used to generate *Ung1/2*^{-/-} CH12 cells target the same exon, which was deleted in *Ung1/2*-deficient mice. **d**, Western blot analysis for UNG (UNG1 and UNG2) and β-actin in wild-type, *Fam72a*^{-/-}, *Ung1/2*^{-/-}, *Ung1*^{-/-}, *Ung2*^{-/-}, *Fam72a*^{-/-} *Ung1*^{-/-} and *Fam72a*^{-/-} *Ung2*^{-/-} CH12 cells. Blots are representative of 3 experiments. **e**, Flow cytometry analysis of IgA expression in additional

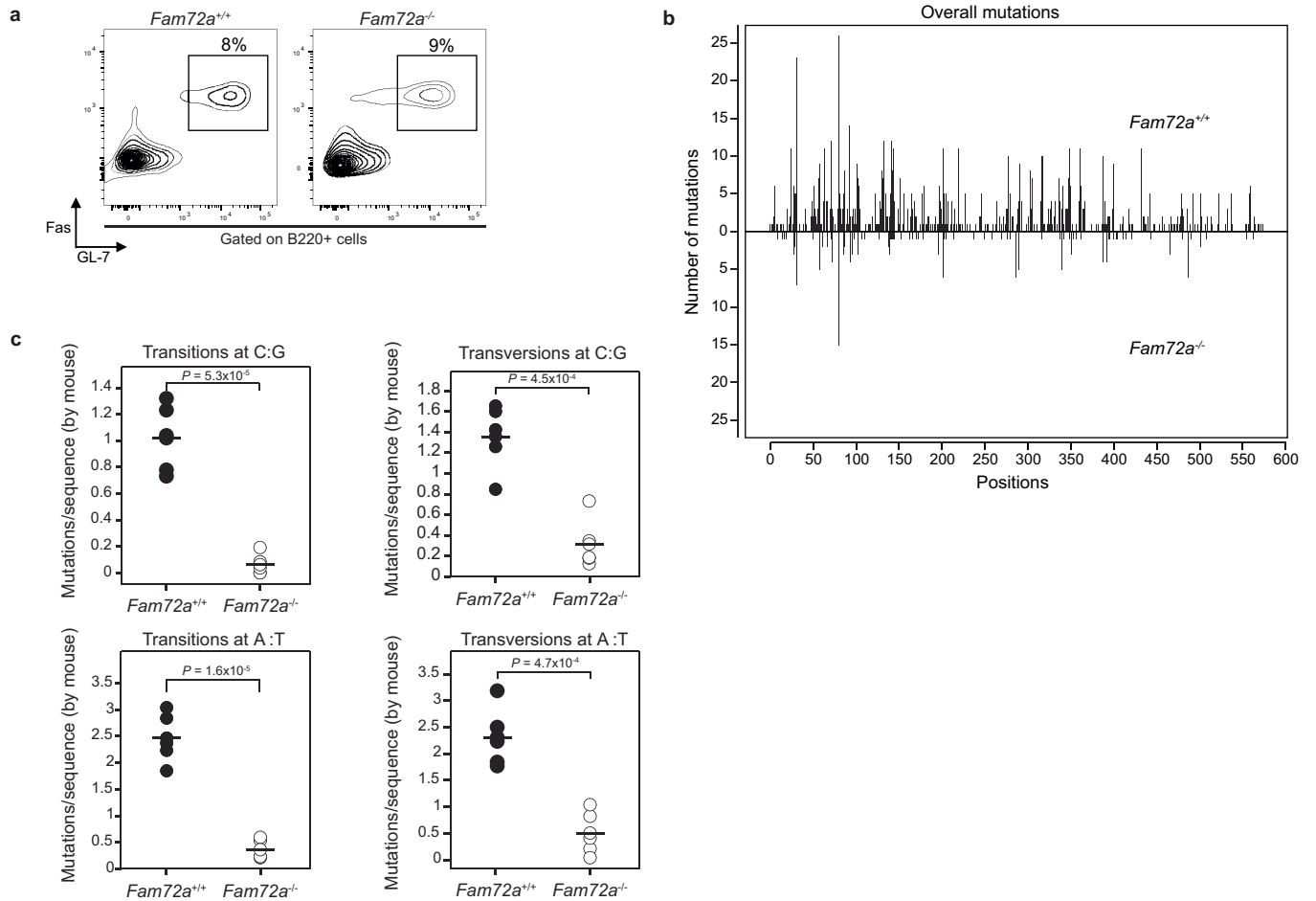
independent wildtype, *Ung1*^{-/-}, *Ung2*^{-/-}, *Fam72a*^{-/-} *Ung1*^{-/-} and *Fam72a*^{-/-} *Ung2*^{-/-} CH12 cells cultured for 72 h with TGFβ, IL-4 and anti-CD40 antibody. **f**, Western blot analysis for UNG (UNG1 and UNG2), NBS1 and Histone H3 on nuclear and chromatin fractions prepared from CH12 cells (pCH12) and *Fam72a*^{-/-} CH12 cells expressing FAM72A or FAM72A^{W125R}. Representative blots of 2 experiments. **g**, RT-qPCR analysis for *Ung1*, *Ung2* and *Igβ* in *Fam72a*^{+/+} and *Fam72a*^{-/-} splenic B cells cultured for 4 days with LPS, IL-4 and anti-IgD-Dextran. Triplicates were normalized to the abundance of *Igβ* and are expressed relative to *Fam72a*^{+/+} B cells, set as 1. Data are presented as mean of five mice per genotype ± s.d. Statistical analysis was performed using two-tailed Student's *t*-test. **h**, Western blot analysis for UNG (UNG1 and UNG2) and β-Actin in wild-type and *Fam72a*^{-/-} CH12 cells cultured in the presence of cycloheximide (CHX) and MG132. **i**, Quantification of the protein level of UNG2, relative to time point zero from 3 independent experiments. Data are presented as mean ± s.d. Statistical analysis was performed using two-tailed Student's *t* test. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 5 | See next page for caption.

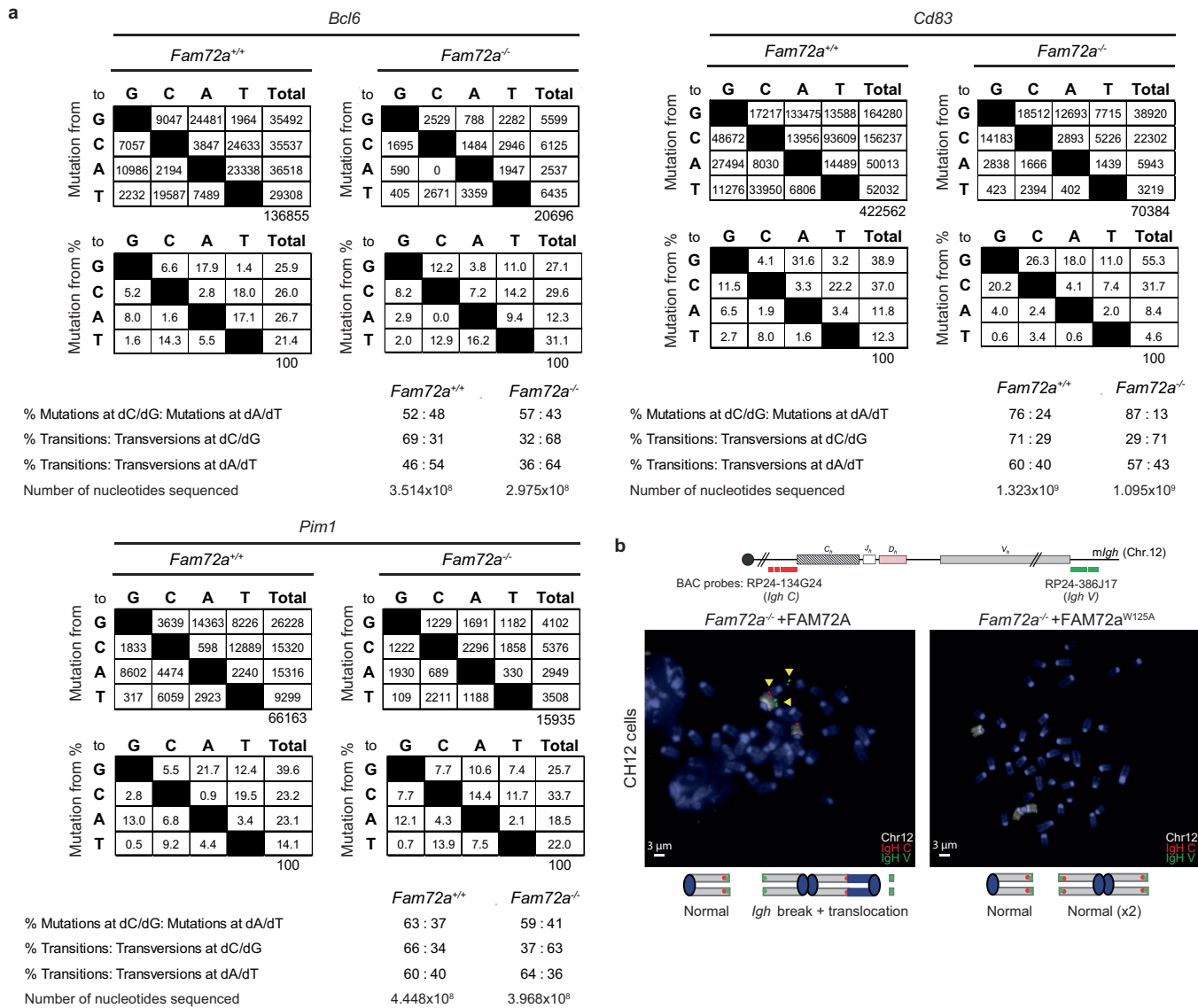
Extended Data Fig. 5 | UNG2 protein levels correlate with the efficiency of CSR. **a**, Western blot analysis for UNG (UNG1 and UNG2) and β -Actin protein levels in wild-type CH12 cells and primary B cells transduced with a retrovirus expressing FAM72A, FAM72A^{W125A} or UNG2 and cultured with TGF β , IL-4 and anti-CD40 antibody or LPS + IL-4 + anti-IgD-Dextran, respectively. **b**, Western blot analysis for UNG (UNG1 and UNG2) and β -actin in *Fam72a*^{-/-} CH12 cells transduced with a retrovirus expressing HIV1-VPR (pMX-VPR), FAM72A and/or FAM72A^{W125A}. **c**, Flow cytometry analysis of IgA expression in *Fam72a*^{-/-} CH12 cells transduced with a retrovirus expressing HIV1-VPR (pMX-VPR), FAM72A and/or FAM72A^{W125A}. Representative plots of three experiments are shown and quantified on the right. P values were determined using two-tailed Student's *t*-test; see Statistics and Reproducibility section in Methods. n.s.: not significant;

<0.01; *<0.001. Horizontal line: mean values. **d**, Flow cytometry analysis of IgA expression in wild-type, *Fam72a*^{-/-} and *Ung1/2*^{-/-} CH12 cells transduced (or not) with a retrovirus expressing UNG2^{Degron} and TIR1 and cultured for 72 h with TGF β , IL-4 and anti-CD40 antibody in the presence or absence of auxin (IAA). Quantification is shown on the right. P-values were determined using two-tailed Student's *t*-test; see Statistics and Reproducibility section in Methods. *<0.05; **<0.005; ***<0.0005. Horizontal line: mean values. Data are from three experiments. **e**, Western blot analysis for UNG (UNG1 and UNG2) and β -Actin in wild-type, *Fam72a*^{-/-}, and *Ung1/2*^{-/-}, transduced (or not) with a retrovirus expressing UNG2^{Degron} and Tir1 and cultured in the presence or absence of auxin (IAA). Blots **a**, **b** and **e** are representative of 3 independent experiments. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 6 | Somatic hypermutation at the J_H4 intron (J_H4i). a, Flow cytometry analysis of germinal center B cells (B220+ Fas+ GL-7+) isolated from the Peyer's patches of unimmunized *Fam72a^{+/+}* and *Fam72a^{-/-}* mice. Plots are gated on B220+ cells. **b,** Distribution of mutations at the J_H4 intron (J_H4i) in *Fam72a^{+/+}* (top) and *Fam72a^{-/-}* (bottom) sequences obtained from germinal center B cells (B220+ Fas+ GL-7+) isolated from the Peyer's patches (PPs) of

unimmunized *Fam72a^{+/+}* and *Fam72a^{-/-}* mice. **c,** Mean frequencies (at JH4i) of transition and transversion mutations at G/C and A/T base pairs, per sequence for each individual mouse analyzed. Data come from 5 mice of each genotype. p-value was determined using two-tailed Student's *t*-test. Horizontal line: mean values.



Pim1

| | | G | C | A | T | Total |
|---------------|------|------|-------|-------|-------|-------|
| Mutation from | to G | 3639 | 14363 | 8226 | 26228 | |
| | to C | 1833 | 598 | 12889 | 15320 | |
| | to A | 8602 | 4474 | 2240 | 15316 | |
| | to T | 317 | 6059 | 2923 | 9299 | |
| | | | | | | 66163 |

| | | G | C | A | T | Total |
|-----------------|------|------|------|------|------|-------|
| Mutation from % | to G | 5.5 | 21.7 | 12.4 | 39.6 | |
| | to C | 2.8 | 0.9 | 19.5 | 23.2 | |
| | to A | 13.0 | 6.8 | 3.4 | 23.1 | |
| | to T | 0.5 | 9.2 | 4.4 | 14.1 | |
| | | | | | | 100 |

Fam72a^{-/-}

| | | G | C | A | T | Total |
|---------------|------|------|------|------|------|-------|
| Mutation from | to G | 1229 | 1691 | 1182 | 4102 | |
| | to C | 1222 | 2296 | 1858 | 5376 | |
| | to A | 1930 | 689 | 330 | 2949 | |
| | to T | 109 | 2211 | 1188 | 3508 | |
| | | | | | | 15935 |

| | | G | C | A | T | Total |
|-----------------|------|------|------|------|------|-------|
| Mutation from % | to G | 7.7 | 10.6 | 7.4 | 25.7 | |
| | to C | 7.7 | 14.4 | 11.7 | 33.7 | |
| | to A | 12.1 | 4.3 | 2.1 | 18.5 | |
| | to T | 0.7 | 13.9 | 7.5 | 22.0 | |
| | | | | | | 100 |

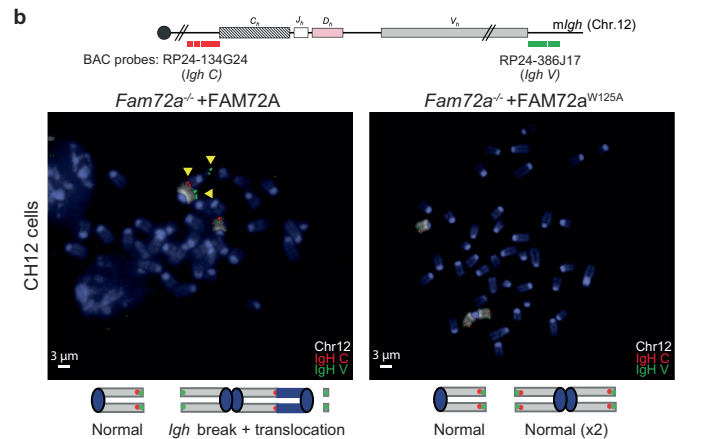
% Mutations at dC/dG: Mutations at dA/dT

% Transitions: Transversions at dC/dG

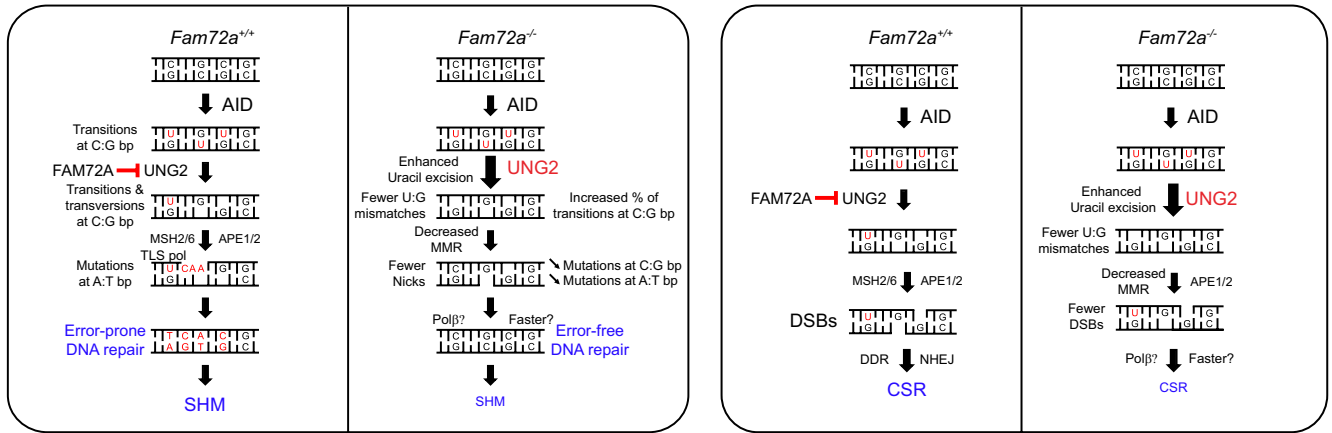
% Transitions: Transversions at dA/dT

Number of nucleotides sequenced

| | | |
|--|------------------------------|------------------------------|
| | <i>Fam72a</i> ^{+/+} | <i>Fam72a</i> ^{-/-} |
| | 63 : 37 | 59 : 41 |
| | 66 : 34 | 37 : 63 |
| | 60 : 40 | 64 : 36 |
| | 4.448x10 ⁸ | 3.968x10 ⁸ |



locus with positions of the BACs used for generation of DNA FISH probes. Lower panel: DNA-FISH on representative metaphases from day 2 stimulated *Fam72a*^{-/-} cells complemented with FAM72A or FAM72A^{W125A} using the *Igh* C BAC probe (red) combined with *Igh* V BAC probe (green) and chromosome 12 paint (white). Yellow arrowheads point to broken or translocated chromosome 12.



Extended Data Fig. 8 | Working model for SHM and CSR in the presence or absence of FAM72A. Immunoglobulin genes are diversified through Somatic Hypermutation (SHM) and Class Switch Recombination (CSR). They are both initiated by the deamination of cytosines in DNA induced by Activation Induced Cytidine Deaminase (AID). The resulting uracils are processed, mainly by the nuclear isoform of Uracil DNA Glycosylase (UNG2), and by proteins of the Base Excision Repair (BER) and Mismatch Repair (MMR) pathways to introduce mutations or double-stranded DNA breaks (DSBs) during SHM (left panel) or

CSR (right panel). FAM72A interacts with UNG2 to control its physiological level by triggering its degradation and to enforce error-prone DNA repair. Consequently, deficiency in *Fam72a* leads to the specific upregulation of UNG2 and its accumulation on chromatin. This would enhance uracil excision, resulting in a reduction in the efficiency of SHM and CSR and enforced error-free DNA repair. Therefore, FAM72A controls the balance between error-prone and error-free DNA repair during antibody diversification.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | | |
|-----|-----------|
| n/a | Confirmed |
|-----|-----------|
- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
 - A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
 - The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
 - A description of all covariates tested
 - A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
 - A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
 - For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
 - For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
 - For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
 - Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

- | | |
|-----------------|--|
| Data collection | No software was used. |
| Data analysis | gRNA enrichment analysis was performed with the MAGeCK-0.5.7 software (https://sourceforge.net/p/mageck/wiki/Home/) and PoolQ-2.2.0 (https://portals.broadinstitute.org/gpp/public/software/poolq). SHM analysis was done with SHMTool (http://shmtool.montefiore.org/cgi-bin/p1). Off-target mutation analysis was done with DeMinER 3.0 (https://github.com/fboyerCRIBL/deminer). |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The datasets generated were submitted to GEO (GSE184145).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|-----------------|--|
| Sample size | No sample-size calculation was performed. For in vitro analysis, experiments were reproduced at least three times with multiple independent clones. For ex vivo experiments, five mice of each genotype were analyzed in at least two independent experiments. |
| Data exclusions | No data were excluded. |
| Replication | Experiments were successfully repeated at least 3 times. |
| Randomization | Non applicable as a small number of samples were handled. |
| Blinding | Non applicable as a small number of samples were handled. |

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| | |
|-------------------------------------|---|
| n/a | Involved in the study |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Antibodies |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Human research participants |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

| | |
|-------------------------------------|--|
| n/a | Involved in the study |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

Antibodies

| | |
|-----------------|--|
| Antibodies used | Flag (M2, Sigma, A8592), IgA-Biotin (Southern Biotech, 1040-08), Ung (Gift from B Kavli, 6103), AID (IGBMC, AID-2E11), b-Actin (Sigma, A1978), IgG1-Biotin (BD Pharmigen, A85-1), IgG2a-Biotin (Biolegend, RMG2A), IgG2b-Biotin (Biolegend, RMG2B1), IgG3-Biotin (BD Pharmigen, R40-82), IgA-PE (Southern Biotech, 1040-09), GL7-PacificBlue (Biolegend), B220-PE-Cy7 (eBiosciences, RA3-6B2), Fas(CD95)-PE (BD Pharmigen, JO2), Streptavidin-PE (Jackson ImmunoResearch, 016-110-084), Histone H3 (Abcam, AB1791), Nbs1 (Gift from A. Nussenzweig), anti-CD40 (eBioscience, HM40-3), anti-chicken IgM-PE (Southern Biotech, 8310-09). |
| Validation | All antibodies were previously validated for their specific applications (flow cytometry and Western blot) and were validated by the manufacturer or the laboratory which generated them. |

Eukaryotic cell lines

Policy information about [cell lines](#)

| | |
|---|---|
| Cell line source(s) | CH12 cells (Gift from T. Honjo), Bosc23 (Gift from M. Nussenzweig), slgM+ ϕ V- AID-/- DT40 cells (Gift from S. Conticello), HEK293T (ATCC). |
| Authentication | All cell lines were previously authenticated for their ability to undergo CSR (CH12/FACS analysis), for producing infectious viral particles (Bosc23 and HEK293T/FACS analysis) and for inducing SHM (slgM+ ϕ V- AID-/- DT40 cells/FACS analysis). |
| Mycoplasma contamination | All cell lines were mycoplasma-free. |
| Commonly misidentified lines (See ICLAC register) | No commonly misidentified cell lines were used. |

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

| | |
|-------------------------|--|
| Laboratory animals | Fam72a ^{-/-} mice were generated by KOMP (C57BL6/N, 8-12 week old males and females). All mice were housed under specific pathogen free conditions. |
| Wild animals | The study did not involve wild animals. |
| Field-collected samples | The study did not involve samples collected from the field. |
| Ethics oversight | Animal work was performed under ethics-approved protocols by the Cometh' ethics committee (APAFIS#23104-2019112915476749 / D672183) |

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

| | |
|---------------------------|---|
| Sample preparation | Cells were stained with appropriate antibodies and stained with DAPI or Topro-3 to exclude dead cells. |
| Instrument | LSRFortessa (BD Biosciences) and FACSAria II (BD Biosciences). |
| Software | FlowJo |
| Cell population abundance | Purity of sorted populations, as determined by post-sort FACS analysis was above 95%. |
| Gating strategy | Dead cells were excluded with DAPI (or Topro-3), then gated on FCS/SSC, then on relevant stainings (EGFP, B220, etc). |

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Étude de l'impact des facteurs induits par l'hypoxie sur l'expression optimale de l'enzyme AID au cours du processus de commutation isotypique des immunoglobulines

Résumé

Les immunoglobulines sont synthétisées par les lymphocytes B. La diversité présente dans le répertoire des cellules B résulte de deux étapes majeures. La première se produit lors du développement des cellules B dans la moelle osseuse, où le mécanisme de recombinaison V(D)J assemble les régions variables VH et VL. La seconde a lieu au cours des réponses immunitaires, grâce à deux mécanismes distincts : l'hypermutation somatique (HMS) et la commutation isotypique (CI). Tant la CI que l'HMS dépendent de l'action de l'enzyme AID (activation-induced cytidine deaminase). Malgré son rôle crucial dans l'immunité humorale, AID présente un potentiel oncogénique significatif en raison de sa capacité à causer des dommages à l'ADN en dehors des loci des immunoglobulines. L'identification des protéines régulatrices d'AID permet une meilleure compréhension de la manière dont cet équilibre est régulé. L'approche de ce projet consiste à identifier les protéines associées à AID par une méthodologie protéomique, en les isolant au moyen de réactions de co-immunoprécipitation et en les identifiant à l'aide de la spectrométrie de masse. Ce travail de recherche a abouti à l'identification de plus de 300 protéines, ainsi qu'à la découverte et à la caractérisation fonctionnelle de quatre complexes protéiques d'importance, jouant un rôle crucial dans la régulation d'AID. Il a été démontré que, lors des réponses immunitaires, des zones d'hypoxie sont observées au sein des centres germinaux (GC), et que la réponse à l'hypoxie module la réponse humorale. Mes recherches ont ainsi mis en lumière que l'expression d'AID et l'efficacité de la recombinaison de classe des immunoglobulines sont conditionnées par le facteur inductible par l'hypoxie (HIF).

Mots clés : AID, commutation isotypique, protéomique, hypoxie, HIF

Résumé en anglais

Immunoglobulins are produced by B lymphocytes, and the diversity found in the B cell repertoire arises from two major stages. The initial stage occurs during the development of B cells in the bone marrow, where the V(D)J recombination mechanism orchestrates the assembly of variable regions VH and VL. The second takes place during immune responses, facilitated by two distinct mechanisms: somatic hypermutation (SHM) and class switch recombination (CSR). Both CSR and SHM depend on the action of the enzyme AID (activation-induced cytidine deaminase). Despite its crucial role in humoral immunity, AID has significant oncogenic potential due to its ability to cause DNA damage outside of immunoglobulin loci. Identifying the regulatory proteins of AID allows a better understanding of how this balance is regulated. The approach of this project involves identifying proteins associated with AID through a proteomic methodology, isolating them through co-immunoprecipitation reactions, and identifying them using mass spectrometry. This research has led to the identification of over 300 proteins and the discovery and functional characterization of four important protein complexes crucial in AID regulation. It has been demonstrated that, during immune responses, hypoxic areas are observed within germinal centers (GC), and the response to hypoxia modulates the humoral response. My research has highlighted that the expression of AID and the efficiency of class switch recombination of immunoglobulins are conditioned by the hypoxia-inducible factor (HIF).

Keywords : AID, class switch recombination, proteomic, hypoxia, HIF