



HAL
open science

Switch Canonique en Cis ou Trans et Recombinaisons Suicides du Locus IgH

Iman Dalloul

► **To cite this version:**

Iman Dalloul. Switch Canonique en Cis ou Trans et Recombinaisons Suicides du Locus IgH. Médecine humaine et pathologie. Université de Limoges, 2018. Français. NNT : 2018LIMO0049 . tel-03027070

HAL Id: tel-03027070

<https://theses.hal.science/tel-03027070>

Submitted on 27 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Limoges

ED 615 - Sciences Biologiques et Santé (SBS)

UMR CNRS 7276 - INSERM 1262 - CRIBL

Contrôle de la Réponse Immune B et des Lymphoproliférations

Thèse pour obtenir le grade de
Docteur de l'Université de Limoges

Discipline : Science Biologie Santé/Spécialité : Immunologie

Présentée et soutenue par

Iman DALLOUL

Le 26 novembre 2018

**Switch Canonique en Cis ou Trans et Recombinaisons Suicides du Locus
IgH**

Thèse dirigée par Jeanne COOK-MOREAU et Michel COGNÉ

JURY :

Rapporteurs

Mme. Claude-Agnès Reynaud, Directeur de Recherche, Institut Necker-Enfants Malades, Paris.
M. Ahmed Amine Khamlichi, Directeur de Recherche, Institut de Pharmacologie et de Biologie Structurale, Toulouse.

Examineurs

Mme. Marion Espéli, Chargée de Recherche, Institut Cochin, Paris.
M. Pierre Bruhns, Directeur de Recherche, Institut Pasteur, Paris.

Directeurs de thèse

Mme. Jeanne Cook-Moreau, maître de conférences, CRIBL, Université de Limoges
M. Michel Cogné, Professeur des Universités, CRIBL, Limoges





Au cours de ces travaux de thèse, j'ai bénéficié pendant les trois premières années d'un financement par la région Limousin. J'ai eu par la suite un contrat doctoral avec l'université de Limoges pour un an pour finaliser les projets.



Laboratoire CRIBL « Contrôle de la Réponse Immune B et des Lymphoproliférations »
UMR CNRS 7276, INSERM 1262
Faculté de Médecine, CBRS : Centre de Biologie et de Recherche en Santé
2, rue du Docteur Marcland
87025 LIMOGES Cedex



Toutes les lettres ne sauraient trouver les mots qu'il faut...

Tous les mots ne sauraient exprimer la gratitude,

L'amour, le respect, la reconnaissance...

Aussi, c'est tout simplement que

Je dédie cette thèse

A la mémoire de mon papa qui nous a laissés très tôt.

Aucune dédicace ne pourrait exprimer mon amour éternel et ma considération pour les sacrifices que tu m'avais consentis pour créer la personne que je suis maintenant.

*Tu me manques toujours, le temps n'effacera jamais le vide que tu as laissé quand tu es parti.
Puisse Dieu t'accueillir dans ses paradis en récompense de tous les sacrifices et les bonnes actions que tu as réalisés pour ta petite fille.*

A ma très chère maman

La plus belle et la plus tendre créature que Dieu a créée sur Terre.

Rien dans ce monde ne peut récompenser les efforts fournis jour et nuit pour mon éducation.

Puisse Dieu, t'accorder bonheur, santé et longue vie.

***« S'il n'y avait pas d'hiver, le printemps ne serait pas si agréable :
si nous ne goûtions pas à l'adversité, la réussite ne serait
pas tant appréciée »***

• Anne Bradstreet

Remerciements

En guise de reconnaissance, je tiens à remercier sincèrement toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce travail.

Je voudrais tout d'abord adresser mes sincères gratitudee à mes deux directeurs de thèse Dr. Jeanne COOK-MOREAU et Pr. Michel COGNE. Merci beaucoup Jeanne ma maman en France pour ton soutien pendant ces quatre années de thèse et pour le temps que tu m'avais consacré pour élaborer ce travail. Je vous remercie sincèrement Michel en tant que chef et directeur de l'UMR CNRS7276 de m'avoir acceptée comme un nouveau membre dans votre équipe et de m'avoir soutenue et encouragée dans les moments durs pendant la réalisation de ma thèse.

Je voudrais ensuite remercier les membres du jury, Claude-Agnès Reynaud, Ahmed Amine Khamlichi, Marion Espéli et Pierre Bruhns qui ont bien voulu accepter d'être les juges de ce travail. C'est un honneur pour moi de discuter mes résultats de recherche à propos desquels vos idées peuvent m'inspirer dans ma future carrière.

Je souhaiterais également remercier tous mes amis aussi membres de l'UMR 7276. Merci beaucoup Batoul et Hussein mes voisins et mes collègues, Nour, Leena, Racha, Valéria, Hend, Jean-Marie, Anne, Victoria, Ophélie, Sébastien et Audrey, vos présences dans le labo m'ont beaucoup encouragé, vous allez me manquer beaucoup, je vous souhaite une bonne continuation et une vie pleine de bonheur et de succès. Merci aussi pour Nivine, Omar, Nicolas et Anderson, je vous souhaite une bonne continuation dans vos futures carrières.

Je voudrais sincèrement remercier François Boyer, sans toi c'était impossible de faire toutes les analyses bio-informatiques. J'ai eu de la chance de travailler avec quelqu'un de si gentil que toi, rendant très difficile de trouver des mots qui peuvent exprimer ma sincère gratitude.

Mes sincères remerciements vont à Jean-Claude et sa femme Cécile. Je ne vais jamais oublier votre gentillesse avec moi.

Pour leurs encouragements et leurs assistances je remercie également Eric, Yves, Sandrine Le Noir, Christophe, Laurent Stéphanie et Anne.

Un grand merci aux ingénieurs du labo Claire, Mylène pour m'aider avec la cytométrie en flux, je ne vais plus vous embêter. Merci aussi à Julie, Cendrine, Christelle, Marie C, Marie M et Zéliha vous allez toutes me manquer.

Merci aussi à Aurore, Christian, Alexis, Brice et Sarah.

Je remercie Sylvie, Bernadette et David mais aussi Emilie, Victor et Valentin.

Je voudrais également remercier Dr. Yolla El Makhour pour ses conseils et de m'avoir encouragée pour parcourir ce long chemin de mes études.

Mes sincères remerciements vont également aux personnes qui m'ont soutenues de loin tous les jours, Zahraa El Bazzal, Najwa Skafi, Rana Bezri et Hussein Kanaan, j'espère que l'amitié qui nous a réunie persiste pour toujours, j'ai de la chance d'avoir des personnes comme vous à côté de moi.

Mes sincères gratitudes vont à ma famille qui m'a encouragé et m'a donné la force pour continuer ma thèse. Merci beaucoup Maman de me donner beaucoup d'amour et de tendresse, je ne vais jamais trouver les mots pour te remercier assez. Merci mon petit frère Mohammad mon petit chouchou ☺, tu étais toujours mon refuge dans mes moments difficiles ; mes grandes sœurs Chadia et Katia vous m'avez toujours soutenu pour arriver jusqu'ici, votre petite sœur est docteur maintenant, je sais que vous êtes fières de moi « Love you so much ♥ ». Un grand merci pour ma sœur et ma collègue dans le bureau et les manip Zeinab aussi membre de l'UMR7276. Ta présence dans le labo et ton soutien m'ont beaucoup aidé pour dépasser les moments durs de la vie que nous avons vécus ensemble. Merci pour ta cuisine qui ne m'a jamais faire sentir que je suis loin de la maison. J'adresse mes sincères remerciements à mon ange gardien mon papa qui m'a quitté au début de ma thèse. Tu m'avais beaucoup soutenu pendant toutes mes études. Ce travail est le fruit de tes sacrifices que tu as consentis pour mon éducation. J'ai voulu vraiment que tu seras avec moi dans ce jour aussi important comme tu m'avais promis juste avant de commencer ma thèse pour te rendre fier de moi. Je ne pourrais jamais te remercier assez.

A mes parents

A mes sœurs Chadia, Katia et Zeinab

A mon frère Mohammad

Contexte Bibliographique

Résumé

L'activation des cellules B est connue pour s'accompagner de remodelages des gènes d'immunoglobulines qui permettent la maturation d'affinité des régions variables d'Ig par hypermutation somatique SHM et la commutation de classe CSR. Ces deux processus sont sous le contrôle de la région régulatrice 3' (3'RR) du locus IgH. Pendant la CSR, le locus IgH subit des changements tridimensionnels mettant les régions switch ciblées par AID à proximité de la région 3'RR afin de faciliter la recombinaison. La sous-unité MED1 du complexe Médiateur favorise cette interaction à longue distance avec la 3'RR mais elle intervient aussi dans la transcription germinale qui précède la CSR afin de faciliter l'activité d'AID. Comme récemment démontré chez la souris, la région 3'RR peut aussi être ciblée par des recombinaison médiée par AID, mais contrairement à la CSR, ce type de recombinaison qui joint la région S μ et la 3'RR et qui s'appelle recombinaison suicide du locus IgH ou LSR entraîne une délétion complète de l'ensemble des gènes constants conduisant à la mort des cellules B par la perte de l'expression du BCR. Nous montrons maintenant que la LSR induite par AID se produit également dans les cellules B humaines activées avec les deux régions 3'RR (3'RR1 en aval de C α 1 et 3'RR2 en aval de C α 2) et elle touche l'allèle fonctionnel. Elle peut aussi être bi-allélique marquée par l'absence de ce type de recombinaison dans les plasmocytes de la moelle mais aussi dans les cellules B mémoires au repos du sang. La LSR se produit à un niveau élevé lorsque les cellules B mémoires sont réactivées. Toutes nos conditions de stimulation utilisées *in-vitro* induit la LSR, sans permettre de discerner comment se fait « le choix » entre la CSR et la LSR. Nos résultats montrent par contre que la sous-unité MED1 semble influencer la transcription de la 3'RR et la recombinaison LSR chez la souris. L'inactivation conditionnelle de MED1 influence l'accessibilité transcriptionnelle et donc les recombinaisons sans affecter les marques épigénétiques du locus IgH. Cette étude de MED1 a aussi révélé que l'ensemble des processus stimulés par l'IgH 3'RR sont « Médiateur-dépendants » (SHM, CSR sans distinction de la cis et la trans-CSR, expression augmentée du locus dans les plasmocytes...), comme semble l'être également le processus de choix des segments variables au cours des réarrangements V_HDJ_H.

Mots clés : 3'RR, locus IgH, CSR, LSR, complexe Médiateur.

Abstract

B-cell activation is accompanied by remodeling of immunoglobulin genes resulting in affinity maturation of Ig variable regions by somatic hypermutation (SHM) and class switch recombination (CSR). These two processes are under the control of the 3' regulatory region (3'RR) of the IgH locus. During CSR, the IgH locus undergoes three dimensional changes bringing the AID-targeted switch regions near the 3'RR region to facilitate recombination. The MED1 subunit of the Mediator complex promotes this long-distance interaction with the 3'RR, but it is also implicated in germinal transcription preceding CSR in order to facilitate AID activity. As recently demonstrated in mice, the 3'RR region can also be targeted by AID-mediated recombination, but unlike CSR, this type of recombination joining the S μ region and 3'RR (called Locus Suicide Recombination or LSR) results in a complete deletion of all the constant genes leading to B-cell death by loss of B Cell Receptor expression. We now show that AID-mediated LSR also occurs in activated human B cells with the two 3'RR (3'RR1 downstream of C α 1 and 3'RR2 downstream of C α 2) and affects the functional allele. It can also be bi-allelic marked by the absence of this type of recombination in plasma cells of the bone marrow but also in quiescent blood memory B cells. LSR occurs at high level when the memory B cells are reactivated. All in-vitro stimulations induce LSR, without identifying conditions favoring either CSR and the LSR. Our results also show that the MED1 subunit appears to influence 3' RR transcription and LSR in mice. Conditional inactivation of MED1 influences transcriptional accessibility and therefore recombination without affecting epigenetic markers of the IgH locus. This study also revealed that all the processes controlled by the 3'RR are "mediator-dependent" (SHM, CSR without distinction between *Cis* and *Trans*-CSR, increased expression of the IgH locus in the plasma cells ...), as well as the choice of variable segments during V_HDJ_H rearrangements.

Keywords: 3'RR, IgH locus, CSR, LSR Mediator complex

Table des matières

Remerciements	6
Résumé	10
Abstract	11
Table des matières	12
Table des illustrations.....	17
Introduction Générale.....	18
Chapitre I : La lymphopoïèse B.....	22
I. Les immunoglobulines et leur loci	22
1. Les loci des chaînes légères.....	22
a. Le locus Kappa.....	22
b. Le locus Lambda.....	22
2. Le locus de la chaînes lourde IgH	23
3. Obtention d'une immunoglobuline fonctionnelle.....	25
II. Moelle osseuse : Phase précoce de la lymphopoïèse B	26
1. Cellules souches hématopoïétiques	27
2. Stade pré-pro-B	27
3. Stade pro-B.....	27
4. Stade pré-B	29
5. Stade B immature	31
III. Phases tardives, antigène dépendante, du développement B.....	32
1. Signalisation du BCR	32
2. Stades des cellules B transitionnelles	34
3. Stade des cellules B matures	35
a. Les cellules B1.....	35
b. Les cellules B2.....	36
i. Les cellules B de la zone marginale MZ.....	36
ii. Les cellules folliculaires FO et formation du centre germinatif.....	37
c. Remaniements secondaires des loci d'Ig induit par AID au sein des centres germinatifs.....	39
i. Hypermutation somatique ou SHM.....	40
ii. Commutation de classe ou CSR.....	43

• Réparation des cassures doubles brin par la voie non homologues NHEJ.....	45
• Réparation des cassures doubles brins par la voie alternative A-EJ.....	46
• Recombinaison <i>trans</i> ou <i>trans</i> -Switch : cas particulier de CSR.....	47
4. Stade de différenciation finale : Sécrétion des anticorps et mémoire immunitaire	48
IV. Activateurs transcriptionnels du locus IgH.....	49
1. L'activateur intronique E μ	49
2. La région régulatrice 3'RR.....	49
a. Structure de la 3'RR.....	49
b. Activité des enhanceurs de la 3'RR lors de l'ontogénie B et les remaniements qui touchent le locus IgH.....	52
c. Implication de la 3'RR dans la structure tridimensionnelle du locus IgH.....	54
d. La région 3'RR: cible pour la recombinaison suicide du locus IgH ou LSR.....	55
Chapitre II : Complexe Médiateur au niveau du locus IgH.....	58
1. Structure du complexe Médiateur	58
2. Fonctions du complexe Médiateur	59
a. Implication du complexe Médiateur dans la transcription	60
b. Répression de la transcription par le Médiateur	61
c. Formation d'une boucle durant la transcription	62
d. Rôle du Médiateur dans les cellules B.....	63
3. Implication du complexe Médiateur en pathologie	65
Objectifs et problématiques.....	68
Article 1 : Article soumis dans Plos Research et accepté dans BioRxiv	71
Article 2: Article soumis dans <i>The Journal of Experimental Medicine</i>	73
Article 3 : Lettre acceptée dans Cellular and Molecular Immunology.....	75
Discussion et perspectives.....	76
Annexe 1: (Article).....	82
Annexe 2: (Article).....	83
Annexe 3: (Revue).....	84
Références bibliographiques	85

Abréviations

3'RR: 3' Regulatory Region

53pB1: p53-binding protein 1

α : alpha

β : bêta

δ : delta

ε : epsilon

γ : gamma

μ : mu

ψ: pseudogène

κ : kappa

λ : lambda

A

Ac: Anticorps

A-EJ: Alternative End Joining

Ag: Antigène

AID: Activation-Induced-Cytidine-Deaminase

APOBEC: Apolipoprotein B Editing Catalytic Polypeptide

ARN pol II: ARN Polymérase II

ATM: Ataxia Telangiectasia Mutated

B

Bcl-XL: B-cell lymphoma-extra large

BCR: B Cell Receptor

BER: Base-Excision Repair

BLIMP1: B Lymphocyte-Induced Maturation Protein 1

BTK: Bruton tyrosine kinase

C

C : Région Constante (CL : région constante de chaînes légères/CH : région constante de chaînes lourdes)

Ca2+: Calcium

CTD: Carboxyl-Terminal-Domain

CD: Cluster of Differentiation

CDK8: Cyclin-Dependent Kinase 8

cEμ: core Eμ

CHA1: Catabolic (L-serine) (L-threonine) deaminase

CHIP: Chromatin Immunoprecipitation

CLP: Common Lymphoid Precursors

CMH: Complexe Majeur d'Histocompatibilité

CSR: Class Switch Recombination

CTCF: CCCTC-binding factor

CtIP: C-terminal-binding protein (CtBP)-Interacting Protein

D

D : segment de Diversité (DL : région de diversité de chaînes légères/DH : région de diversité de chaînes lourdes)

dA: désoxy-Adénosine monophosphate

dC: désoxy-Cytidine

dG: désoxy-Guanosine

dT: désoxy-Thymidine

dU : désoxy-Uridine

DAG : Diacylglycerol

DDX1 : DEAD-Box Helicase 1

DNA-PKcs : DNA-Protein Kinase catalytic subunit

DNase I: DésoxyriboNucléase I

DSB: Double Strand Break

E

E2A: E box binding protein 2A

EBF: Early B Cell Factor

EXO1: DNA Exonuclease 1

F

FDC: Follicular Dendritic Cells

FO: Follicular Zone

FOXO1: Forkhead box protein O1

G

G4: G-quadruplex

GC: Germinal Center

GSK3: Glycogène Synthase Kinase 3

GTFs: General Transcription Factors

H

H: Heavy (lourde)

H2AX: H2A histone family member X

H3/H4: Histone 3/Histone 4

HIGM: Hyper-IgM syndrome

hs : hypersensible à la DNase I

HSCs: Hematopoietic Stem Cells

I

I: Intronic promotor

Ig: Immunoglobulin

IgH: Immunoglobulin Heavy chain

IgL: Immunoglobulin Light chain

IL: Interleukin

ILR: Interleukin Receptor

INO1: Inositol-3-phosphate synthase

IP3: Inositol 1,4,5-Trisphosphate

IR: Inverted Repeat

IRF4: Interferon Regulatory Factor 4

ITAM: Immunoreceptor Tyrosine-based Activation Motif

J

J : segment de Jonction (JL : de chaînes légères/JH : de chaînes lourdes)

K

Kb: Kilo base

L

L : Leader ou Light (légère)

LB: Lymphocyte B

LMPP: Lymphoid-primed MultiPotent Progenitor

LPS: Lipopolysaccharide

LS: Like-Switch

LSR: Locus Suicide recombination

LT: Lymphocyte T

M

MAPK: Mitogen-Activated Protein Kinase

MALT: Mucosae Associated Lymphoid Tissue

MAR: Matrix Attachment Region

MCL-1: Megalencephalic Leukoencephalopathy with Subcortical Cysts 1

MDC1: Mediator of DNA damage checkpoint protein 1

MED: Sous-unité du complexe Médiateur

MLH1: MutL Homologue 1

MMR: MisMatch Repair

MPP: Multipotential Early Progenitors

MRN: Mre11 Rad50

Nbs1 complex

Msh2 et Msh6: Mut S homolog 2 and 6

mTORC2: mammalian target of rapamycin complex 2

MZ: Marginal Zone

N

NBS1: Nijmegen Breakage Syndrome 1
NES: Nuclear Export Signal
NFAT: Nuclear Factor of Activated T cells
NF-κB: Nuclear Factor-kappa B
NHEJ: Non-Homologous End Joining
NK: Natural Killer
NLS: Nuclear Localization Signal

P

Pax5: Paired-box gene 5
Pb: Paire de base
PI3K: Inositol 1,4,5-trisphosphate 3-kinases
PIC: Pre-Initiation Complex
PIP2: Phosphatidylinositol-4,5-bisphosphate
PIP3: Phosphatidylinositol-3,4,5-trisphosphate
PKC: Protéine Kinase C
PLCγ2: Phospholipase Cγ2
PMS2: PostMeiotic Segregation 2
pol II: ARN polymerase II
PU.1: PUrine box factor-1

R

RAG1/2: Recombination Activating Gene 1 or 2
RSS: Recombination Signal Sequence

S

S: Région Switch
SCID: Severe Combined Immunodeficiency
SHM: Somatic Hypermutation
SL: Surrogate Light Chain
SMUG1: Single-strand selective monofunctional uracil DNA glycosylase
SSB: Single Strand Break
SYK: Spleen tyrosine kinase

T

TACI: Transmembrane Activator and CAML Interactor
TCR: T Cell Receptor
TFII: Transcription Factor II
Tfh: T follicular helper
TFR: Follicular Regulatory T cells
TGFβ: Transforming Growth factor bêta
T_H: T Helper
TLR: Toll Like Receptor

U

UNG: Uracil DNA Glycosylase

V

V : région Variable et segment de Variabilité (VL : de chaînes légères/ VH : de chaînes lourdes)

X

XBPI: X-box Binding Protein 1
XRCC4: X-Ray Repair Cross Complementing 4

Z

Zn²⁺: Zinc

Table des illustrations

Figure 1: Structure d'une immunoglobuline : Exemple d'une IgG1	21
Figure 2: Représentation schématique de l'organisation des loci des chaînes légères d'immunoglobuline chez l'Homme et chez la souris	23
Figure 3: Représentation schématique de l'organisation du locus des chaînes lourdes d'immunoglobuline chez l'Homme et chez la souris	24
Figure 4: Développement des lymphocytes B (exemple chez la souris)	26
Figure 5: Recombinaison VDJ et VJ	28
Figure 6: Développement du pré-BCR des cellules B	30
Figure 7: Voies de transduction du signal induites par le BCR	34
Figure 8: Développement des cellules B transitionnelles et matures	35
Figure 9: AID et modulation des réactions au sein des centres germinatifs des organes lymphoïdes	38
Figure 10: Structure de l'AID et désamination de l'ADN	40
Figure 11: Hypermutation somatique	41
Figure 12: Réparation des lésions induites par AID	42
Figure 13: Représentation schématique de la CSR et de la structure de la boucle R	44
Figure 14: Réparation des cassures par les voies NHEJ et A-EJ	46
Figure 15: Structure de la 3'RR chez l'Homme et chez la souris	51
Figure 16: Structure tridimensionnelle du locus IgH facilitant la CSR dans les cellules B matures	55
Figure 17: Localisations des LS dans la région 3'RR murin et humain	56
Figure 18: Mécanisme de recombinaison LSR	57
Figure 19: Représentation schématique du complexe Médiateur humain	59
Figure 20: Activation de la transcription par le complexe Médiateur	61
Figure 21: Inhibition de la transcription par le module CDK8	62
Figure 22: Formation de la boucle d'ADN par le complexe Médiateur	63
Figure 23: Récapitulatif des fonctions du complexe Médiateur	66

Introduction Générale

Le système immunitaire est constitué d'un ensemble coordonné d'éléments permettant de lutter contre toute agression extérieure de l'organisme (parasite, virus, bactérie) mais aussi de contrôler certains dérèglements cellulaires endogènes aboutissant à la formation des cellules malignes. Cette protection repose essentiellement sur la capacité du système immunitaire à tolérer le soi et à éliminer le non soi, quand celui-ci se manifeste dans un contexte d'inflammation ou de « danger ». Pour accomplir son rôle, ce système fait intervenir différents organes et cellules.

Notre organisme dispose de deux systèmes de défense contre toute agression : l'immunité **innée** et l'immunité **adaptative**. La première appelée encore immunité **naturelle** constitue la première ligne de défense, elle mobilise notamment les cellules phagocytaires (macrophage, cellules dendritiques et les polynucléaires) ainsi que les cellules tueuses NK (*Natural Killer*) pour éliminer les pathogènes qui ont franchi les barrières physiques de l'organisme (peau, muqueuses...). C'est une réponse immunitaire immédiate et non-spécifique, très sensible aux « signaux de danger ». Un réseau de communication fort existe entre ces deux versants de l'immunité afin d'éliminer le pathogène qui menace notre organisme. Certaines cellules du système immunitaire inné agissent en alertant le système immunitaire **adaptatif** (aussi appelé **acquis**), mobilisant des cellules spécialisées de deux grands types : les lymphocytes B et T qui interviennent lors des réponses immunitaires humorale et cytotoxique respectivement. Cette réponse apparaît plus tardivement et permet de compléter et de renforcer les réponses du système immunitaire innée. Contrairement à l'immunité innée, l'immunité adaptative possède une mémoire immunitaire assurant une réaction plus rapide après une deuxième rencontre avec l'antigène. C'est une réponse immunitaire spécifique à longue durée.

Les lymphocytes B (LB) et les lymphocytes T (LT) sont capables de reconnaître une grande diversité d'antigènes (Ag) via leurs récepteurs, le BCR (*B Cell Receptor ou récepteur des cellules B*) et le TCR (*T Cell Receptor ou récepteur des cellules T*). Chez les lymphocytes B ce récepteur membranaire est composé d'une immunoglobuline (Ig) associée aux deux protéines membranaires Iga ou CD79a et Igβ ou CD79b qui assurent la transduction d'un signal constitutif de survie et de signaux d'activation des cellules B suite à une reconnaissance spécifique d'un antigène. Après activation, les lymphocytes B se différencient en plasmocytes sécrétant des immunoglobulines solubles appelées anticorps Ac et qui ont la même spécificité que le BCR dans le but de neutraliser ou d'éliminer un antigène.

Les immunoglobulines sont des glycoprotéines hétérodimériques constituées de deux chaînes lourdes IgH (*Immunoglobulin Heavy Chain*) et de deux chaînes légères IgL (*Immunoglobulin Light Chain*) reliées entre elles par des ponts disulfures (**Figure 1**). Chaque chaîne d'immunoglobuline est composée d'une région variable et d'une région constante. L'association des régions variables des chaînes lourdes et des chaînes légères forme le site de reconnaissance de l'antigène ; la région constante détermine les fonctions effectrices de l'immunoglobuline telles que la fixation du complément et la liaison sur les récepteurs Fc de plusieurs types cellulaires. On peut distinguer chez l'Homme comme chez la souris deux types de chaînes légères, kappa (κ) et lambda (λ) et cinq classes de chaînes lourdes (μ , δ , γ , ϵ , α) qui déterminent respectivement les cinq classes d'immunoglobulines : IgM, IgD, IgG, IgE et IgA. Il existe plusieurs sous-classes d'IgG chez la souris : IgG1, IgG2a, IgG2b et IgG3 et chez l'Homme : IgG1, IgG2, IgG3 et IgG4. On distingue également chez l'Homme deux sous-classes d'IgA : IgA1 et IgA2.

Les IgM sont les premières immunoglobulines produites lors d'un premier contact avec un antigène. Les IgD sont les deuxièmes immunoglobulines produites et elles sont dans la plupart des cas co-exprimées avec les IgM. Après infection ou immunisation, les lymphocytes B peuvent substituer l'immunoglobuline exprimée à leur surface (IgM ou IgD) vers les autres classes d'Ig par un mécanisme appelé commutation de classe ou CSR (*Class Switch Recombination*) ou switch. La CSR est une recombinaison intra ou inter-chromosomique (alors inter-allélique) qui a lieu sur le locus IgH entre deux régions switch (S) situées en amont de chaque segment constant (C) qui code la région constante des Igs. Elle permet la synthèse de différents isotypes d'Ig caractérisés par des fonctions effectrices distinctes sans affecter la spécificité antigénique.

La commutation de classe est contrôlée par plusieurs éléments *cis*-régulateurs qui ont été décrits pour le locus IgH. Il s'agit de promoteurs, de l'activateur intronique $E\mu$ et de la région régulatrice 3'RR (*3' Regulatory Region*) située en 3' du locus IgH. Chez la souris, la région 3'RR s'étend sur une distance longue de plus de 30 kb et son activité est restreinte à la lignée lymphoïde B. L'organisation de la région 3'RR d'IgH est complexe et inclut plusieurs activateurs transcriptionnels (ou « *enhancers* ») identifiés initialement comme des sites hypersensibles (hs) à la DNase I. Chez l'Homme il existe deux régions 3'RR (3'RR-1 : située en aval du gène $C\alpha 1$ et 3'RR-2 située en aval du gène $C\alpha 2$) contenant chacune trois enhancers : *hs3*, *hs1-2* et *hs4*.

Pendant la CSR, le locus IgH subit des changements tridimensionnels mettant les promoteurs, les activateurs et les régions switch à proximité les uns les autres pour faciliter le switch. Récemment, Thomas-Claudepierre et ses collègues ont montré que l'inactivation de la sous unité MED1 du complexe médiateur (régulateur transcriptionnel agissant sur les interactions à longue distance) diminue cette interaction au niveau du locus IgH, ce qui se traduit par une réduction de switch associée à une baisse de la transcription de la région switch acceptrice (Thomas-Claudepierre et al., 2016).

Dans notre laboratoire, un nouveau type de recombinaison a été décrit pour la première fois chez la souris : la recombinaison suicide du locus IgH ou LSR (Locus Suicide Recombination) dépendante de l'enzyme Activation-Induced-Cytidine Deaminase (AID) (Péron et al., 2012a). Ce phénomène induit la perte d'expression du BCR membranaire et donc du « signal tonique » constitutivement transmis par ce récepteur, conduisant ainsi à la mort de la cellule B.

Dans un premier temps je détaillerai lors de l'introduction les différentes étapes du développement lymphocytaire B, en expliquant plus particulièrement les phénomènes de recombinaisons géniques du locus IgH. Je décrirai ensuite l'implication de certaines protéines qui peuvent altérer la commutation de classe des immunoglobulines en insistant surtout sur le complexe Mediator.

La partie expérimentale est divisée en deux parties. Dans la première partie, je me suis particulièrement attachée à mettre en évidence le phénomène de LSR chez l'Homme *in-vivo* mais aussi *in-vitro* après stimulation et d'essayer de décrypter son rôle dans la physiologie et la pathologie du lymphocyte B. La description des recombinaisons LSR, précédemment chez la souris et maintenant chez l'Homme, nous a conduits ensuite à rechercher une implication éventuelle du phénomène de LSR dans les lymphopénies B acquises et sévères participant significativement à la gravité du tableau d'immunodéficience globale et prolongée dans certaines maladies comme le sepsis.

La deuxième partie concerne le rôle du complexe Médiateur dans la commutation de classe par l'étude de ce phénomène chez les souris spécifiquement déficientes dans la lignée lymphoïde B pour la sous unité MED1. Les effets sur la sécrétion des Ig, les différentes sous populations des lymphocytes B, l'hypermutation hypersomatique, la CSR et LSR, la transcription et la fréquence respective des événements de *cis* versus *trans* switch ont été examinés.

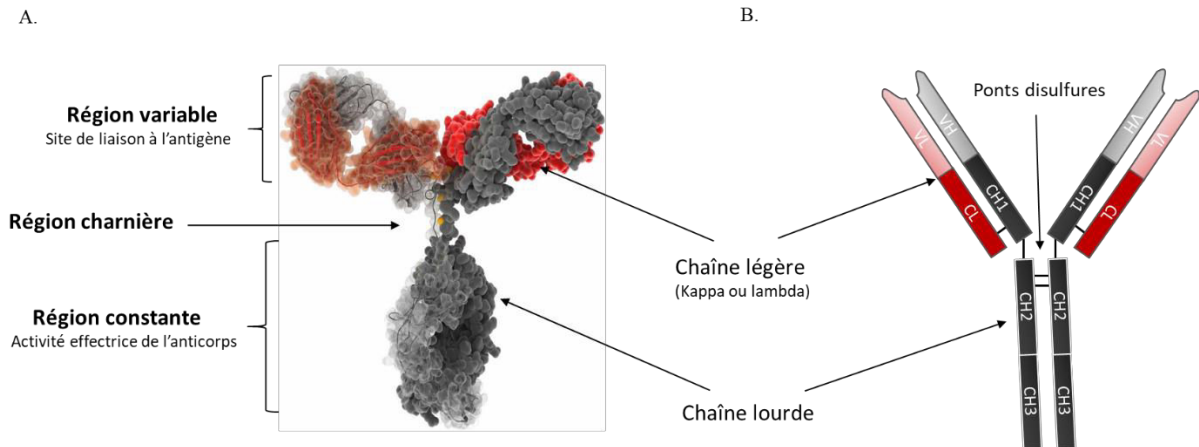


Figure 1: Structure d'une immunoglobuline : Exemple d'une IgG1

A. Représentation tridimensionnelle d'une immunoglobuline

B. Représentation schématique d'une immunoglobuline. Elle est constituée de deux chaînes lourdes représentées en noir et de deux chaînes légères représentées en rouge. Les chaînes lourdes sont constituées d'un domaine variable VH et de 3 domaines constants (CH1, CH2 et CH3). Les chaînes légères sont constituées aussi d'un domaine variable VL et d'un domaine constant CL. La région charnière située entre le domaine CH1 et CH2 confère la flexibilité à l'immunoglobuline.

Chapitre I : La lymphopoïèse B

I. Les immunoglobulines et leur loci

Chez l'Homme comme chez la souris la synthèse des immunoglobulines est assurée par des gènes regroupés sur 3 loci distincts. Les 2 loci $Ig\kappa$ et $Ig\lambda$ codent les chaînes légères de l'immunoglobuline alors que le locus IgH code les chaînes lourdes. Chaque locus est formé de segments variables (V), de segments de diversité (D : uniquement sur le locus IgH), des segments de Jonction (J) et enfin les segments constants (C). Les segments V(D)J codent la partie variable de l'anticorps qui fixe l'antigène alors que les segments constants codent la partie constante qui déterminent l'activité effectrice d'une Ig (Early et al., 2004; Giudicelli et al., 2005).

Comme j'ai travaillé durant ma thèse sur l'Homme et sur la souris qui ont une organisation très similaire des gènes d'Ig, je détaillerai plus loin leurs loci tout en gardant dans l'esprit les variations qui peuvent exister entre les différentes souches de souris (Rauta et al., 2012).

1. Les loci des chaînes légères

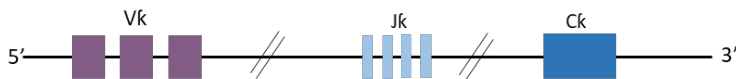
a. Le locus kappa

Le locus kappa est localisé sur le chromosome 2 chez l'Homme (chromosome 6 chez la souris). Il comporte 76 gènes $V\kappa$ chez l'Homme dont 31 à 35 sont fonctionnels (174 gènes $V\kappa$ chez la souris mais seulement 94 à 96 sont fonctionnels), 5 gènes $J\kappa$ (chez la souris 4 gènes sont fonctionnels) et 1 seul gène $C\kappa$. Ces gènes codent les chaînes légères de l'immunoglobuline (**Figure 2A**) (Giudicelli et al., 2005).

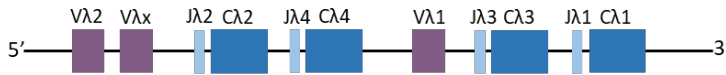
b. Le locus lambda

Le locus lambda est localisé sur le chromosome 22 chez l'Homme (chromosome 16 chez la souris). Ce locus a une organisation particulière, il comporte chez l'Homme 69 gènes $V\lambda$ mais seulement 31 sont fonctionnels et 7 régions des gènes $J\lambda-C\lambda$. Les clusters $J-C\lambda 1$, $J-C\lambda 2$, $J-C\lambda 3$, et $J-C\lambda 7$ sont fonctionnels alors que les clusters $J-C\lambda 4$, $J-C\lambda 5$ et $J-C\lambda 6$ sont non fonctionnels (**Figure 2C**). Le locus $Ig\lambda$ murin comporte 3 gènes $V\lambda$ organisé en 2 segments $V2-Vx-J2-C2-J4-C4$ et $V1-J3-C3-J1-C1$ (**Figure 2B**) (van der Burg et al., 2001; Giudicelli et al., 2005; Tümkaya et al., 2001).

A. Locus Ig κ (chromosome 6 pour la souris et 2 pour l'Homme)



B. Locus Ig λ murin (chromosome 16)



C. Locus Ig λ humain (chromosome 22)

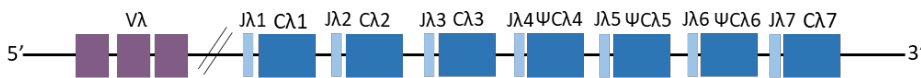


Figure 2: Représentation schématique de l'organisation des loci des chaînes légères d'immunoglobuline chez l'Homme et chez la souris

Les segments variables V sont représentés en violet, les segments de jonction J sont représentés en bleu clair et les segments constants C en bleu foncé.

A. Représentation du locus Ig κ situé sur le chromosome 6 chez la souris et 2 chez l'Homme. Ce locus a la même organisation chez les deux espèces mais avec un nombre différent de segments V, J et C.

B. Représentation schématique du locus Ig λ chez la souris située sur le chromosome 16.

C. Représentation schématique du locus Ig λ chez l'Homme situé sur le chromosome 22.

2. Le locus de la chaînes lourde IgH

Le locus IgH est situé sur le chromosome 14 chez l'Homme (chromosome 12 chez la souris). Du télomère vers le centromère, il comporte 45 à 51 gènes V_H fonctionnels (103 gènes V_H chez la souris), 25 gènes de diversité D_H fonctionnels (9 gènes D_H chez la souris), qui n'existent pas sur les loci des chaînes légères et 6 gènes J_H fonctionnels (4 gènes J_H chez la souris) (Collins and Jackson, 2018). Ce locus est le « théâtre » du sujet principal de ma thèse ; j'en détaillerai plus loin l'organisation tout en soulignant les différences qui existent entre l'Homme et la souris.

Au nombre de 9 chez l'Homme et 8 chez la souris, les gènes C_H codent les régions constantes des différentes classes et sous-classes des immunoglobulines : C μ , C δ , C γ_3 , C γ_1 , C α_1 , C γ_2 , C γ_4 , C ϵ , C α_2 mais aussi 2 pseudogènes C $\psi\gamma$ et C $\psi\epsilon$ chez l'Homme et C μ , C δ , C γ_3 , C γ_1 , C γ_2b , C γ_2a , C ϵ , C α chez la souris. Ce cluster des gènes constants existe en une seule copie chez la plupart des espèces comme chez la souris. D'autres espèces comme l'Homme et les singes de l'ancien monde possèdent une duplication de ce cluster de gènes constants. Cette duplication ancestrale s'est produite il y a plus que 60 millions années chez l'ancêtre commun (hominoïdes)

donnant lieu à 2 régions de gènes constants du locus IgH chez l'Homme, les gorilles et les chimpanzés (D'Addabbo et al., 2011; Pinaud et al., 2011). Chaque gène constant C_H contient plusieurs exons qui codent les différents domaines de l'Ig (CH1, CH2, CH3) pour l'IgA, IgD, IgG. L'IgE et l'IgM possèdent un domaine CH4 supplémentaire. La production de la forme sécrétée d'une Ig sous forme d'un anticorps est assurée par la présence d'une région spécifique située à la fin du dernier exon constant (CH3 ou CH4), et conservée ou non selon un mécanisme d'épissage alternatif du messenger. Alternativement, la forme membranaire est codée par intégration en 3' du messenger d'un ou plusieurs exons supplémentaires qui codent les parties membranaire et intracytoplasmique. Toutes les Igs à l'exception d'IgE et d'IgM possèdent une région « charnière » entre leurs deux domaines CH1 et CH2 conférant leur flexibilité (**Figure 3**) (Schroeder and Cavacini, 2010).

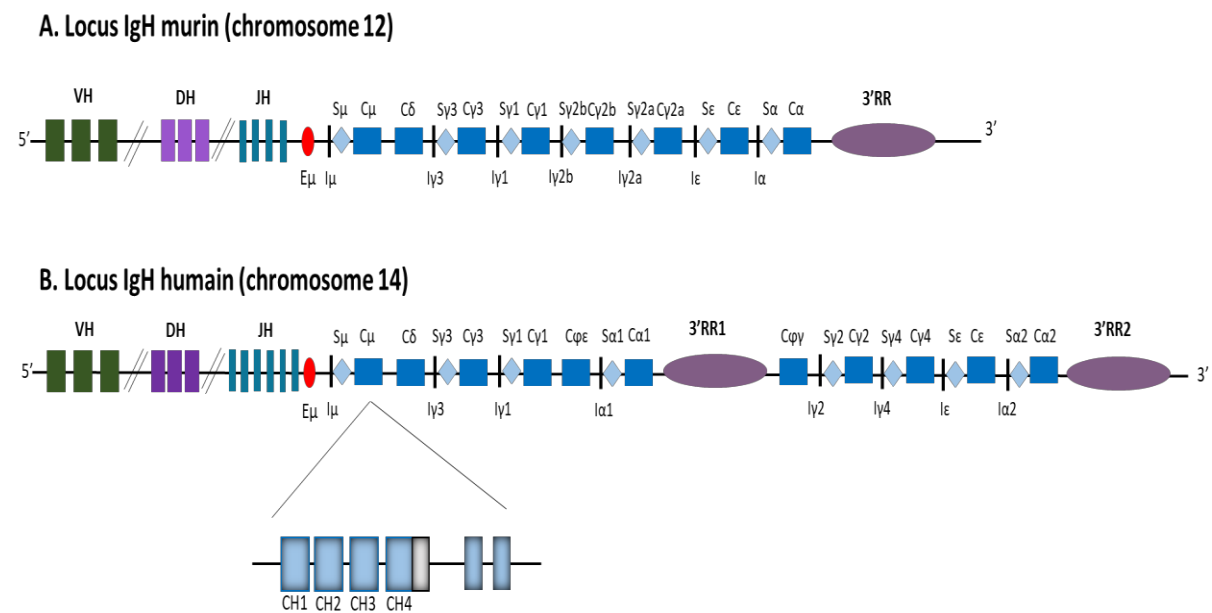


Figure 3: Représentation schématique de l'organisation du locus des chaînes lourdes d'immunoglobuline chez l'Homme et chez la souris

Pour les loci des chaînes lourdes, les segments variables V_H sont représentés en vert, les segments de jonction D_H sont représentés en violet foncé et les segments de jonction J_H en turquoise. Les segments constants C représentés en bleu foncé sont précédés par des régions switch S représentées par des losanges en bleu clair. Un promoteur germinale « I » est situé du côté 5' de chaque région switch. Chaque segment constant contient plusieurs exons CH (CH1, CH2, CH3 et un CH4 supplémentaire pour C_μ et C_ϵ). L'activateur E_μ situé en amont des gènes constants est représenté par un rond rouge.

A. Représentation du locus IgH murin situé sur le chromosome 12 possédant une seule région 3'RR.

B. Représentation du locus IgH humain situé sur le chromosome 14 possédant deux régions 3'RR (3'RR1 et 3'RR2).

En amont de chaque gène CH (à l'exception de C δ qui est précédé d'une région non répétitive $\Sigma\delta$) est localisée une région switch (S) dont la taille varie entre 1 à 12 kb. Ces régions contiennent des motifs répétés riches en GC (chez les mammifères), tels que TGGGG, GGGGT, GGGCT, GAGCT et AGCT constituant les cibles principales de l'enzyme AID qui initie la commutation de classe ou CSR (Shinkura et al., 2003; Zarrin et al., 2004). La délétion de ces motifs répétés dans la région S μ chez la souris entraîne une diminution drastique de la CSR (Khamlichi et al., 2004). Ces régions switch sont également précédées du côté 5' par des promoteurs germinaux d'initiation de la transcription appelé les promoteurs I_H (I μ , I γ , I ϵ et I α) qui permettent l'initiation de la transcription qui progresse ainsi à travers la région switch puis la région constante pour générer le transcrit germinale I α -S α -CH α nécessaire pour le switch (Shinkura et al., 2003).

Le locus IgH contient également 2 activateurs transcriptionnels majeurs appelés E μ et la région régulatrice 3' ou 3'RR qui régulent les recombinaisons et l'expression des gènes de ce locus.

3. Obtention d'une immunoglobuline fonctionnelle

Dans leur configuration germinale, les gènes d'Ig sont non fonctionnels. Ils doivent alors subir plusieurs remaniements et recombinaisons génétiques de leurs loci pour générer une Ig fonctionnelle. Ces événements ont lieu durant la lymphopoïèse B. On distingue des événements dits « précoces » qui ont lieu dans la moelle osseuse et d'autres dits « tardifs » qui ont lieu dans les organes lymphoïdes secondaires (rate, plaques de Peyer, amygdales, ganglions lymphatiques) (Pieper et al., 2013).

Les événements précoces précèdent la rencontre avec un antigène, ils consistent à induire des recombinaisons V_H(D)J_H du locus des chaînes lourdes et V_LD_LJ_L pour les loci des chaînes légères afin de générer le répertoire des lymphocytes B capables de connaître un grand nombre d'antigènes (Jung et al., 2006; Pieper et al., 2013; Schlissel, 2004).

Les événements tardifs regroupent la CSR et la SHM (*Somatic hypermutation*) qui suivent la rencontre avec un antigène. La CSR touche le locus IgH et permet de changer la classe d'Ig d'une IgM vers une autre classe. C'est un mécanisme de recombinaison entre les gènes codant les domaines constants, diversifiant ainsi ses activités effectrices sans toucher la spécificité contre l'antigène. La SHM introduit des mutations à fréquence élevée dans les régions qui codent les parties variables de l'immunoglobuline dans le but de modifier le site de fixation de

l'antigène. Cette étape est suivie par une sélection des cellules produisant des anticorps à haute affinité dans les centres germinatifs (Shinkura et al., 2004).

II. Moelle osseuse : Phase précoce de la lymphopoïèse B

Chez l'Homme comme chez les autres mammifères, les lymphocytes B se développent chez l'adulte dans la moelle osseuse à partir des précurseurs hématopoïétiques. Les cellules B de la moelle osseuse, après avoir subi plusieurs remaniements au niveau de leurs loci IgH et IgL, expriment à leur surface un BCR de classe IgM, puis elles quittent la moelle osseuse et migrent vers les organes lymphoïdes secondaires pour se différencier en cellules B matures (Pieper et al., 2013). Les différents stades de développement B ainsi que les marqueurs exprimés à leurs surfaces sont représentés dans la **Figure 4**.

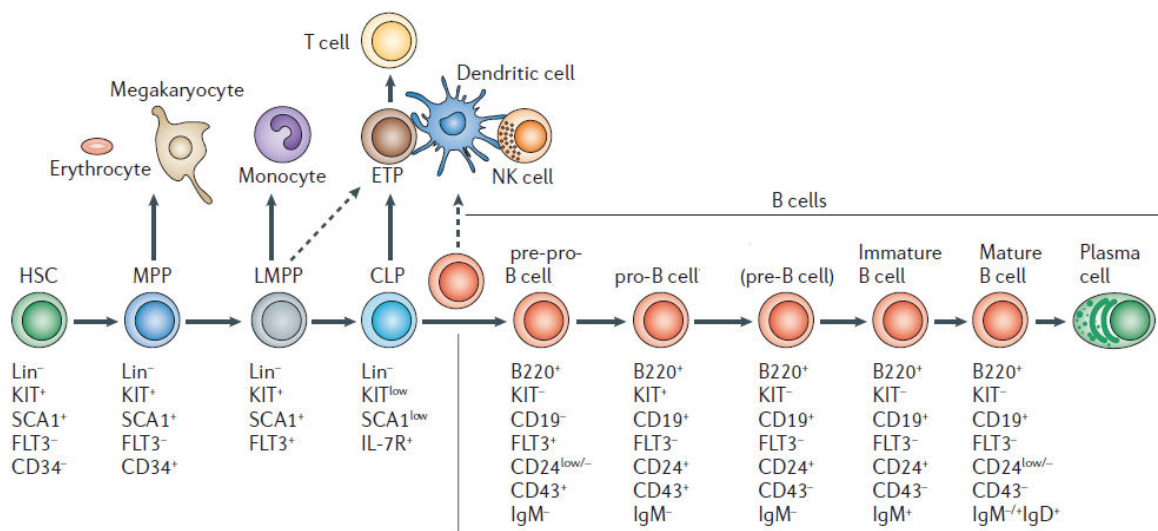


Figure 4: Développement des lymphocytes B (exemple chez la souris)

Les cellules souches hématopoïétiques HSC sont censées générer les progéniteurs hématopoïétiques multipotents (MPP). Les progéniteurs lymphoïdes multipotents (LMPP) sont les cellules qui ont perdu leur potentiel de différenciation érythro-mégacaryocytaire mais qui conservent le potentiel de se différencier en cellules myéloïdes et en lymphocytes B et T. Les cellules qui peuvent donner naissance à des cellules B et T mais pas à des cellules de la lignée myéloïde constitue les progéniteurs lymphoïdes communs (CLP). Les précurseurs de cellules B négatifs pour l'immunoglobuline de surface cellulaire (Ig) mais positifs pour le marqueur B220 peuvent être divisés au cours de développement en quatre stades successifs de différenciation, marqués par leur expression différentielle de certains marqueurs de surface (cellules pré-pro-B, cellules pro-B et cellules pré-B). Les cellules B immatures générées à partir des pré-B sortent de la moelle osseuse et atteignent la rate, où elles mûrissent en cellules B matures périphériques et en plasmocytes (*Adaptée de la revue: Nagasawa, 2006*).
ETP, progéniteur précoce de la lignée cellulaire T ; NK, Natural killer.

1. Cellules souches hématopoïétiques

Les cellules souches hématopoïétiques ou HSCs (*Hematopoietic Stem Cells*) sont des cellules multipotentes rares. Elles sont capables d'une part de s'autorenouveler (maintenant constante la réserve de ces cellules), mais aussi de se différencier en progéniteurs des différentes lignées cellulaires sanguines (Riether et al., 2015).

Dans un premier temps, les cellules hématopoïétiques sont capables de se différencier en cellules progénitrices multipotentes précoces MPP (*Multipotential Early Progenitors*) pouvant se différencier en progéniteurs myéloïdes et lymphoïdes précoces sous l'action de PU.1, Gfil et Ikaros. Des niveaux faibles de PU.1 favorisent la différenciation lymphoïde alors que des niveaux élevés favorisent la différenciation myéloïde. L'expression de Gfil dans les MPP inhibe l'expression de PU.1, alors favorisant indirectement la lymphopoïèse B. Ikaros antagonise PU.1 en sur-exprimant Gfil. Sous l'action d'Ikaros, E2A et Ebf-1, les progéniteurs lymphoïdes se différencient pour donner les précurseurs lymphoïdes communs CLP (*Common Lymphoid Precursors*) qui génèrent les lignées cellulaires B, T et NK. L'engagement dans la lignée B se fait sous l'influence de plusieurs facteurs qui favorisent la différenciation lymphocytaire B tout en réprimant les lignées T et NK comme E2A, Ebf1, Pax5 et FOXO1 (Pelayo et al., 2012; Zhang et al., 2013).

2. Stade pré-pro-B

Les cellules pré-pro-B représentent les précurseurs les plus précoces. Elles sont caractérisées par l'expression de certains marqueurs de surface tels que B220, CD79 (Ig α et Ig β) et CD43 mais elles n'expriment pas le CD19 (Hardy et al., 2007; Nagasawa, 2006). La signalisation par le récepteur IL-7R dans ces cellules joue un rôle dans la transition vers le stade pro-B dans lesquelles se dérouleront les recombinaisons V_HDJ_H du locus IgH (Clark et al., 2014).

3. Stade pro-B

A ce stade les précurseurs pro-B expriment les marqueurs CD19 à leurs surfaces (Hardy et al., 2007). Et c'est à ce stade que s'initient les premiers réarrangements qui codent les chaînes lourdes des immunoglobulines entre les segments D_H et J_H. Presque toutes les cellules pro-B ont un réarrangement DJ_H sur les deux allèles. Ce réarrangement marque la transition vers le stade pré-B au cours duquel un réarrangement fonctionnel et productif est réussi entre un segment V_H et le segment DJ_H (Ehlich et al., 1994).

La recombinaison V(D)J et VJ commencent par l'action de deux protéines RAG1 et RAG2 (*Recombination Activating Gene 1 or 2*) qui reconnaissent et clivent des séquences signaux de recombinaison (RSS, *Recombination Signal Sequences*). Ces RSS sont formées par une séquence heptamère (CACAGTG) et une séquence nonamère (ACAAAAACC) hautement conservées séparées par des espaces non conservés de 12 (12-RSS) ou 23 (23-RSS) pb (paires de base) (Shetty and Schatz, 2015). Les recombinaisons surviennent sur les RSS contenant des espaces différents 12-RSS et 23-RSS suivant la règle 12/23. Cette règle évite la recombinaison entre un segment V_H avec un segment J_H mais aussi la recombinaison entre deux segments V_H ou D_H . Si le réarrangement V_H-DJ_H est productif, la cellule présente alors un réarrangement V_H-DJ_H sur un allèle et DJ_H sur l'autre allèle. Si ce réarrangement est non productif, un autre réarrangement V_H-DJ_H aura lieu sur le deuxième allèle. Dans ce cas-là le réarrangement peut être productif et la cellule continue son développement mais il peut aussi être non productif et la cellule meurt à cause de la perte de l'expression d'une Ig à sa surface. Un réarrangement sur trois (1/3) est fonctionnel et permet la synthèse d'une chaîne lourde fonctionnelle (2/3 sont non fonctionnels) (**Figure 5**) (Jung et al., 2006).

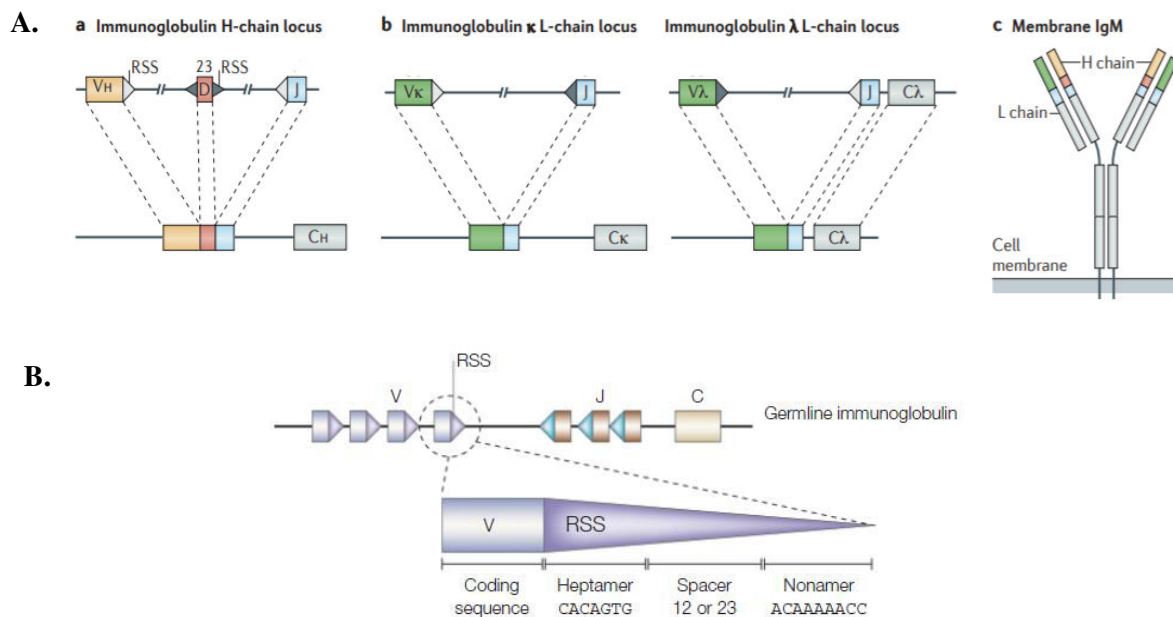


Figure 5: Recombinaison VDJ et VJ

A. Assemblage par recombinaison d'ADN entre des segments de gènes d'Ig. Les recombinaisons s'effectuent entre les segments V_H , D et J_H sur le locus IgH (a), entre les segments V_κ et J_κ pour le locus Ig κ et entre V_λ et J_λ pour le locus Ig λ (b) afin d'exprimer un BCR de type IgM à la surface des cellules B (c). Les triangles représentent le signal de recombinaison RSS. Les deux mécanismes de recombinaison VDJ et VJ s'effectuent de la même façon (règle 12/23) (*Extrait de la revue: Nemazee, 2017*). **B.** Structure des RSS. Les RSS sont formés d'un héptamère et d'un nonamère séparés avec un espaceur de 12 ou 23 pb (*Adaptée de la revue: Roth, 2003*).

4. Stade pré-B

Durant ce stade ces cellules expriment le marqueur CD25. C'est un stade crucial qui amène à l'expression d'un BCR à la surface des lymphocytes B immatures (Mårtensson et al., 2010; Tumang et al., 2011).

Le réarrangement du segment V_H avec le segment DJ_H permet l'expression d'une chaîne μ qui s'associe à une pseudo chaîne légère (*SL, Surrogate Light Chain*) formé par les protéines VpréB (Ig V-like sequence) et $\lambda 5$ (Ig C-like sequence). Seules les chaînes qui peuvent s'associer avec la chaîne SL sont exportées à la surface des cellules pré-B après leur association avec l'hétéromère CD79a et CD79b responsable de la transduction du signal. La signalisation à partir de ce récepteur induit la prolifération et la différenciation des précurseurs pré-B (Abdelrasoul et al., 2018; Mourcin et al., 2012).

On distingue 2 populations de précurseurs pré-B : les grandes cellules pré-B (large pre-B ou early pre-B) et les petites cellules pré-B (small pre-B ou late pre-B) (Cooper et al., 2006). Les grandes cellules pré-B après avoir subi 2 à 5 cycles de division se différencient en petites cellules pré-B non prolifératives (Zhang et al., 2004). A ce stade les petites pré-B expriment à nouveau RAG1 et RAG2 qui remplaceront l'expression du pré-BCR par celle d'un BCR en initiant le réarrangement V-J séquentiel des loci de la chaîne légère d'immunoglobuline (Ig κ puis Ig λ). L'assemblage de la chaîne légère (κ ou λ) avec la chaîne lourde déjà réarrangée permet l'expression d'un BCR de type IgM à la membrane des cellules définies comme des lymphocytes B immatures (**Figure 6**) (Mårtensson et al., 2010).

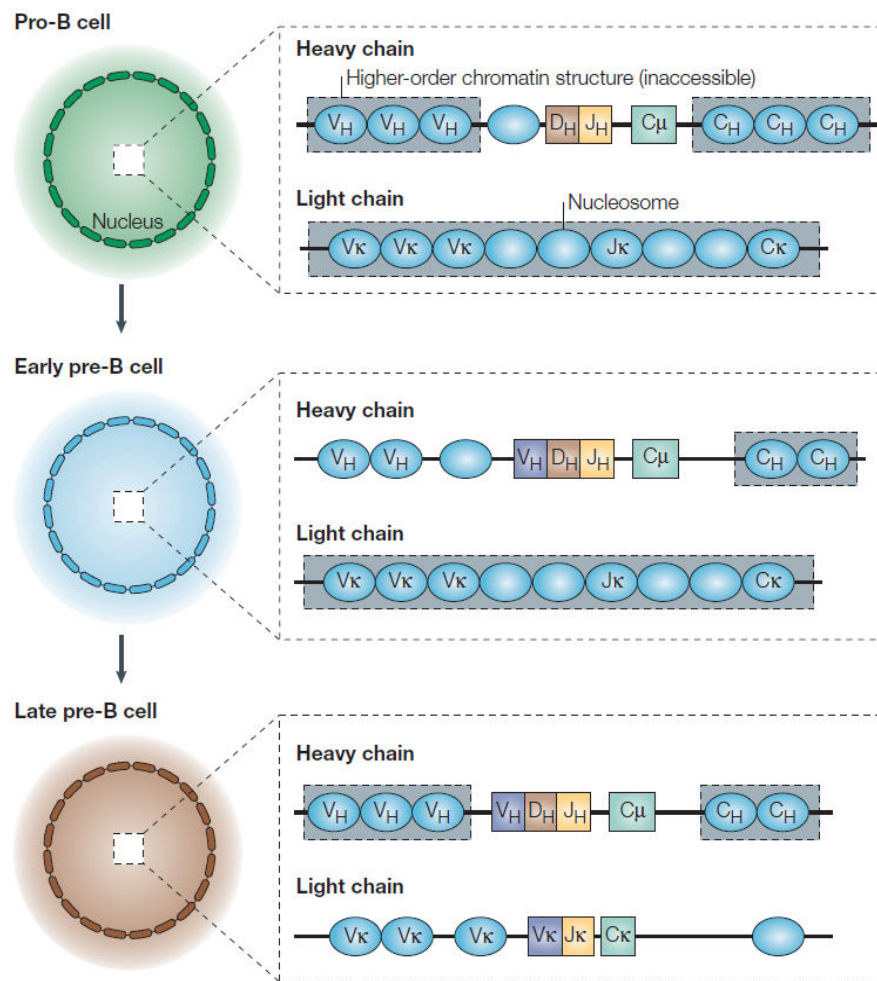
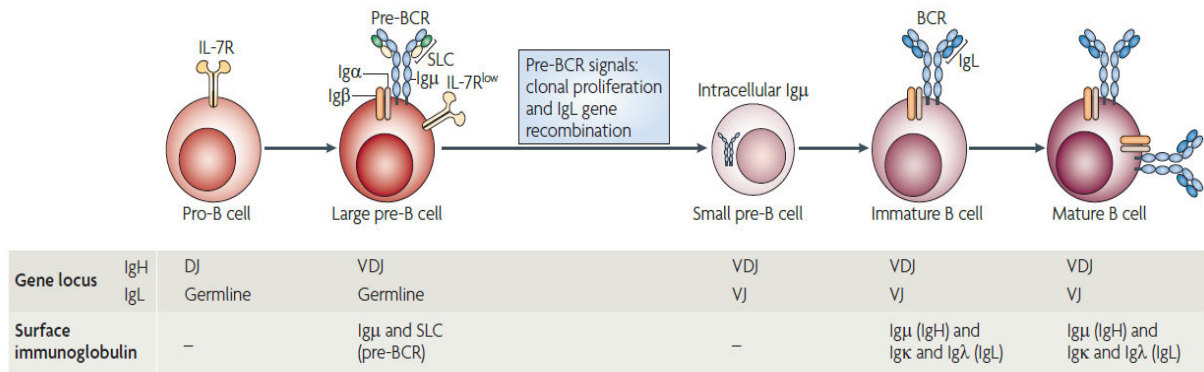


Figure 6: Développement du pré-BCR des cellules B

Le réarrangement du locus de la chaîne lourde IgH est initié au stade pro-B (recombinaison DJ_H) suivi par autre réarrangement V_H-DJ_H au stade pré-B précoce générant ainsi une chaîne Igμ exprimée sous forme d'un pré-BCR à la surface des progéniteurs après assemblage avec une chaîne SL (λ5 et VpréB). Le réarrangement de la chaîne légère dans les petites cellules pré-B aboutit à l'expression d'un BCR avec deux chaînes légères et deux chaînes lourdes de type IgM au stade des cellules B immatures (*Extraite des revues : Herzog et al., 2009 ; Bergman et al., 2004*).

5. Stade B immature

Une fois le BCR présent à la surface des cellules B immatures, il devient finalement capable d'interagir avec plusieurs ligands ; une sélection à ce stade est essentielle afin de tester l'interaction récepteur-ligand. Le contact des cellules immatures avec leur micro-environnement et avec les cellules stromales de la moelle osseuse permet une sélection négative entraînant l'élimination des cellules qui montrent dès ce stade des interactions fortes avec des antigènes du soi. Cette induction de tolérance peut suivre plusieurs voies, de la délétion clonale à l'anergie ou à l'édition du BCR par réinduction de RAG, délétion du segments $V\kappa J\kappa$ réarrangé et apparition d'une nouvelle chaîne légère $V\lambda J\lambda$ modifiant la spécificité antigénique. Seules les cellules qui ne reconnaissent pas ou faiblement les antigènes du soi peuvent quitter la moelle osseuse et migrer vers les organes lymphoïdes secondaires où elles seront testées pour la deuxième fois pour leurs auto-réactivités contre les antigènes du soi avant d'atteindre la phase des lymphocytes B matures. Une auto-réactivité/polyréactivité de faible affinité affecte vraisemblablement les cellules B au compartiment B1, expliquant ultérieurement la production d'anticorps naturels par ce compartiment (Halverson et al., 2004; Hoffman et al., 2016; Nemazee, 2017; Pelayo et al., 2012).

On distingue trois mécanismes de tolérance : l'édition du BCR, la sélection négative et l'anergie (Pelayo et al., 2012).

L'édition du BCR se produit fréquemment chez l'Homme et chez la souris jouant ainsi un rôle crucial pendant le développement du répertoire du BCR (Bergman and Cedar, 2004). L'interaction d'un auto-antigène avec le BCR diminue son expression à la membrane suivie par une augmentation de l'expression de RAG1 et RAG2 afin d'initier le réarrangement $V\kappa J\kappa$ du deuxième allèle du locus $Ig\kappa$ ou bien le réarrangement du locus $Ig\lambda$. En parallèle à la production d'un nouveau réarrangement fonctionnel, le réarrangement préalable de l'un ou des deux allèles kappa pourra être complètement éliminé grâce à un type particulier de réarrangement entre des segments non codants porteurs de RSS, les « κ deleting elements » (κ DE) (Feddersen et al., 1990; Rassenti et al., 1991; Siminovitch et al., 1985). Les cellules qui ne réussissent pas à éditer leur BCR seront éliminées par apoptose ou inactivées par anergie (Halverson et al., 2004 ; Nemazee, 2017).

III. Phases tardives, antigène dépendante, du développement B

Après avoir quitté la moelle osseuse, les lymphocytes B immatures (naïfs) empruntent la circulation sanguine pour atteindre les organes lymphoïdes secondaires (rate, ganglions lymphatiques et les tissus lymphoïdes associés à la muqueuse) où elles rencontreront l'antigène (phase dépendante de l'Ag). Dans ces organes les cellules B peuvent rencontrer l'antigène et alors poursuivre leur développement en se différenciant en plasmocytes sécréteurs d'Ig solubles appelées anticorps (Ac) et/ou en cellules B mémoires (Durandy et al., 2013; Neumann et al., 2015). Durant cette seconde phase de différenciation antigène-dépendante, les lymphocytes B sont capables de produire des anticorps plus affins (grâce au phénomène de SHM) et dont les propriétés effectrices peuvent varier (grâce à la CSR) (Hwang et al., 2015). La signalisation du récepteur des cellules B joue un rôle crucial dans la détermination du devenir des cellules B pendant leur développement tardif lors des réponses immunes. La perturbation d'un composant de la signalisation du BCR peut induire un arrêt de leur maturation. Les progrès récents dans les études des mécanismes moléculaires des voies de signalisation ont aidé à clarifier comment ces signaux régulent la prolifération, la survie et l'apoptose des cellules B le long de leur développement (Niironen and Clark, 2002).

1. Signalisation du BCR

Le BCR a deux rôles principaux. Le premier est de transmettre des signaux qui régulent le devenir des cellules B. Le second est de permettre d'internaliser l'antigène afin de le dégrader puis le présenter aux cellules T CD4, sous forme associée au CMH de classe II, ce qui permet une activation complète dans la phase effectrice (Niironen and Clark, 2002).

Les cellules B normales reçoivent deux types de signaux de leurs BCR. Le premier - appelé tonique - se produit sans une liaison d'un antigène et assure la survie des cellules B circulants et au repos. L'élimination du BCR d'une cellule B entraîne l'apoptose de cette cellule, ce qui suggère que la signalisation tonique du BCR est nécessaire pour la survie normale des lymphocytes B. Le deuxième type de signal est généré par la liaison d'un antigène externe sur le BCR, déclenchant l'activation d'un complexe de signalisation qui transmet le signal à l'intérieur de la cellule (Bojarczuk et al., 2015). Cette liaison de l'antigène active la protéine tyrosine kinase LYN. Ensuite, LYN phosphoryle les motifs ITAMs (*Immunoreceptor Tyrosine-based Activation Motif*) situés sur les queues cytoplasmiques de Iga et Igβ, qui, à leur tour recrutent et activent la protéine Kinase SYK (*Spleen tyrosine kinase*) et la protéine tyrosine

Kinase BTK (*Bruton tyrosine kinase*) (Bojarczuk et al., 2015; Hendriks et al., 2014). Ces événements précoces déclenchent la phosphorylation de nombreux substrats en aval nécessaires à l'activation des seconds messagers pour déclencher à la fin la transcription des gènes associés à l'activation des cellules B (Chaturvedi et al., 2011).

PI3K et PLC γ 2 sont deux enzymes effectrices clés qui génèrent des seconds messagers impliqués dans la signalisation. PI3K phosphoryle le phosphatidylinositol-4,5-bisphosphate (PIP2) pour produire le phosphatidylinositol-3,4,5-trisphosphate (PIP3), qui, à son tour, recrute certaines molécules de signalisation activant la kinase AKT. La voie de signalisation induite par PI3K est un candidat principal pour la modulation de la survie des lymphocytes B. Parmi les molécules en aval qui sont activées après l'activation de PI3K, l'AKT présente un intérêt particulier. En général, l'AKT active des régulateurs impliqués dans le cycle cellulaire et la survie des cellules B (comme MCL-1, BCL-xL, la cyclin D2 et MYC) par une inhibition de la glycoène synthase kinase 3 (GSK3) qui normalement fonctionne comme un régulateur négatif du cycle cellulaire mais aussi par une inactivation de FOXO 1/3 (Forkhead Box Subgroup O 1/3) (une surexpression de FOXO1/3 entraîne un arrêt dans la phase G1 du cycle cellulaire). D'autre part AKT active mTORC2 qui à son tour phosphoryle AKT afin d'augmenter son activité (Bojarczuk et al., 2015; Yusuf et al., 2004).

PLC γ 2 utilise le PIP2 - même substrat que PI3K - mais génère de l'inositol-1,4,5-trisphosphate (IP3) et du diacylglycérol (DAG), qui sont nécessaires pour la libération de calcium intracellulaire (Ca²⁺) et l'activation de la protéine kinase C (PKC), respectivement. La signalisation du calcium active le facteur nucléaire NFAT (*Nuclear Factor of Activated T cells*) alors que la signalisation induite par la PKC active le facteur NF- κ B (*Nuclear Factor- κ B*). La signalisation induite par PLC γ 2 joue aussi un rôle crucial dans l'activation de la voie MAPK (*Mitogen-Activated Protein Kinase*) (Bojarczuk et al., 2015).

Ces différentes voies de signalisation induites par le BCR ne sont pas seulement impliquées dans la prolifération et l'activation cellulaire mais aussi dans la migration des cellules B et l'apoptose (**Figure 7**) (Stevenson et al., 2011).

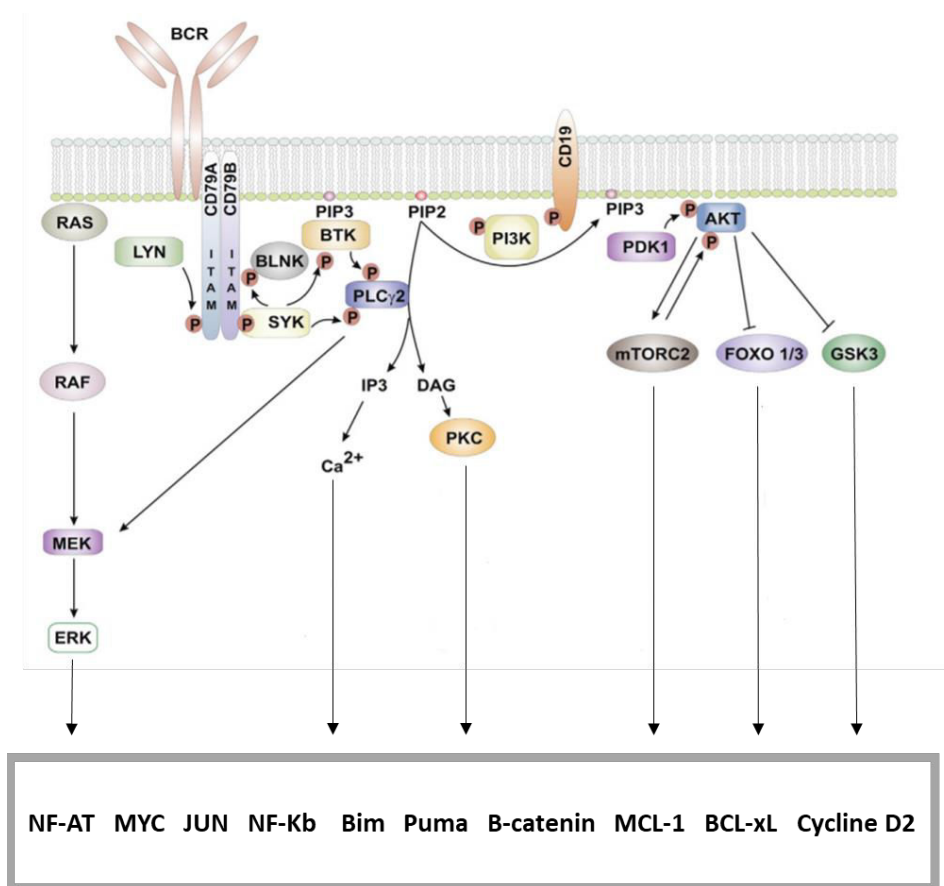


Figure 7: Voies de transduction du signal induites par le BCR

La liaison d'un antigène sur le BCR induit la formation d'un complexe qui déclenche des cascades de signalisation activant plusieurs facteurs impliqués dans la survie, la prolifération, la migration, la différenciation et l'apoptose des cellules B (*Adaptées des revues: Bojarczuk et al., 2015; Steele, 2014*).

2. Stades des cellules B transitionnelles

Il est estimé chez la souris que seulement 10% des cellules B immatures quittent la moelle osseuse pour se diriger vers la périphérie. Ces cellules B transitionnelles constituent un stade intermédiaire entre les cellules B immatures de la moelle osseuse et les cellules B matures (Agrawal et al., 2013; Allman et al., 1993; Teague et al., 2007). En 2001, Alman et ses collègues ont défini 3 sous-populations des cellules B transitionnelles en se basant sur l'expression de l'IgM, CD23 et de la protéine AA4 ou CD93. Les T1 sont AA4⁺IgM^{fort}CD23⁻, les T2 sont AA4⁺IgM^{fort}CD23⁺ alors que les T3 sont AA4⁺IgM^{faible} CD23⁺ (Allman et al., 2001). Les cellules transitionnelles (T1, T2, et T3) sont récemment identifiées chez l'Homme et représentent approximativement 4% des cellules B dans le sang (Marie-Cardine et al., 2008; Martin et al., 2016).

Les cellules T1 et T2 diffèrent par leur réponse à la stimulation du BCR par une induction des signaux différents au niveau de la cellule (Petro et al., 2002) . Les cellules B transitionnelles

auto-réactives sont éliminées par une forte activation du BCR (sélection négative). Les cellules T1 non-réactives développent en cellules T2 qui sont plus résistantes à l'apoptose induit par le BCR, elles se différencient en cellules B folliculaires ou en cellules B de la zone marginale (sélection positive) (**Figure 8**) (Hoek et al., 2006).

Bien que les transitionnelles T3 existent elles ne participent pas à la génération des cellules B matures. Ce sont des cellules auto-réactives qui ne sont pas éliminées par apoptose mais elles sont maintenues dans un état d'anergie (Liubchenko et al., 2012)

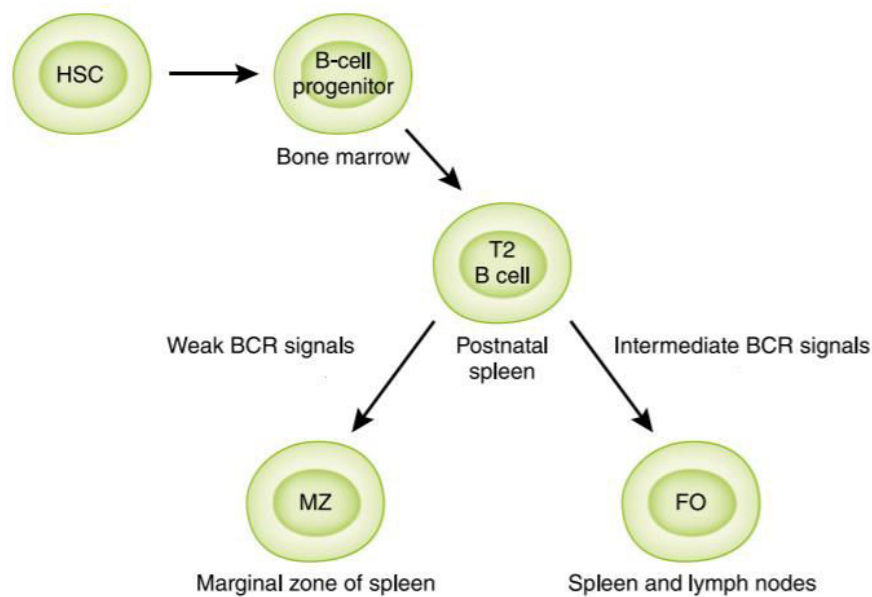


Figure 8: Développement des cellules B transitionnelles et matures

Les cellules B2 se développent à partir de cellules B transitionnelles 2 (T2) dérivées de progéniteurs de cellules B dans la moelle osseuse, avec différenciation subséquente en lignées marginales (MZ) et folliculaires (FO) présentes dans les organes lymphoïdes secondaires. Des signaux plus forts du BCR induisent la différenciation des cellules transitionnelles T2 en cellules B Folliculaires (FO) tandis que des signaux plus faibles du BCR permettent la différenciation en cellules MZ (*Adaptée de la revue: Hoffman et al., 2016*).

3. Stade des cellules B matures

On distingue deux types de cellules B matures qui diffèrent par leur localisation, leur phénotype et leur fonction : les cellules B1 et les cellules B2 (Hampel et al., 2011; Hoffman et al., 2016).

a. Les Cellules B1

Les cellules B1 sont impliquées dans les réponses immunitaires innées sécrétant majoritairement des anticorps de type IgM en absence d'une infection. Ces cellules sont abondantes au niveau des cavités péritonéales et pleurales. Les B1 sont divisées en deux sous-

populations qui diffèrent par l'expression du marqueur CD5 : les B1a (CD19⁺CD5⁺CD11b⁺) et les B1b (CD19⁺CD5⁻CD11b⁺). Les anticorps produits par ces cellules sont fréquemment auto-réactifs et de faibles affinités pouvant se fixer sur les antigènes du soi et déclencher des maladies auto-immunes. D'autre part ces Ac naturels sécrétés par les lymphocytes B1 peuvent aussi jouer un rôle important dans le maintien de l'homéostasie tissulaire en raison de leur capacité à lier les auto-antigènes altérés, tels que ceux exprimés par les cellules apoptotiques. Ces cellules sont générées avant la naissance à partir des progéniteurs au niveau du foie fœtal et qui sont maintenues à l'âge adulte par un auto-renouvellement (Hoffman et al., 2016; Macias-Garcia et al., 2016; Oliveira et al., 2005).

b. Les cellules B2

Les cellules B2 sont divisées en deux populations distinctes : Les cellules B de la zone marginale (*MZ, Marginal Zone B Cells*) et les cellules B folliculaires (*FO, Follicular B Cells*) (Pillai and Cariappa, 2009a). La décision du destin des cellules transitionnelles T2 dépend de la force de la signalisation mais aussi de certains acteurs induits par le BCR. Contrairement aux cellules B folliculaires qui nécessitent une forte activation du BCR, la différenciation des T2 en MZ nécessitent un faible signal du BCR (Pillai and Cariappa, 2009a).

i. Les cellules B de la zone marginale MZ

Ces cellules comme leur nom l'indique sont localisé au niveau de la zone marginale entre la pulpe rouge et la pulpe blanche de la rate. Elles constituent 5% des lymphocytes B de la rate chez la souris (Zouali and Richard, 2011). Ces cellules jouent un rôle vital dans la première ligne de défense contre les infections bactériennes dans le sang en répondant rapidement aux Ags, souvent de manière indépendante des cellules T (Kanayama et al., 2005). Ces cellules peuvent aussi participer dans les réponses de façon T-dépendantes permettant de générer des Ac switchés (Hoffman et al., 2016). Outre une faible activation du BCR, l'engagement des cellules B transitionnelles en cellules B de la zone marginale nécessite également la signalisation de Notch2 exprimé dans les cellules B (Laffleur et al., 2014a). Plusieurs études ont montré un rôle clé de Notch2 dans la génération et le développement des cellules B de la zone marginale dont son inactivation entraîne une réduction dramatique de la productions de ces cellules accompagnée avec une augmentation du nombre des cellules B folliculaire (Saito et al., 2003; Witt et al., 2003).

ii. Les cellules B folliculaires FO et formation du centre germinatif

Les cellules FO représentent la majorité des cellules B de la rate, elles sont comme leur nom l'indique localisées dans les follicules de la rate (Pillai and Cariappa, 2009a). Ces cellules expriment fortement le marqueur CD23 et faiblement le CD21 (Kanayama et al., 2005). Une fois qu'elles atteignent la maturité, les cellules B folliculaires acquièrent la capacité de circuler à travers le sang et la lymphe entre les organes lymphoïdes (Pillai and Cariappa, 2009b).

• Formation du centre germinatif

Suite à une stimulation antigénique de type T-dépendante nécessitant une coopération entre les lymphocytes T et B, les centres germinatifs (GC) sont formés par prolifération des cellules B dans les follicules des tissus lymphoïdes périphériques, y compris la rate, les ganglions lymphatiques, les plaques de Peyer et les amygdales. Au centre de ces follicules, les lymphocytes B sont en contact étroit avec un réseau de cellules dendritiques : les FDC (*Follicular Dendritic Cells*). Ces follicules sont entourés par une zone riche en cellules T appelée « la zone T ». L'activation des cellules B naïves par un antigène entraîne leur migration vers la zone interfolliculaire (frontière entre la zone T et la zone des cellules B) pour interagir avec les cellules T. Ces cellules B activées peuvent se différencier 1) en plasmablastes extrafolliculaires sécréteurs d'anticorps de faible affinité et de courte durée de vie impliqués dans la réponse immédiate contre le pathogène ou 2) participer à la formation du GC afin de générer des cellules B mémoires et des plasmocytes sécréteurs d'anticorps de haute affinité et de longue durée de vie.

Le GC est formé de deux zones distinctes : la zone sombre et la zone claire. Au sein de la zone sombre les cellules B appelées centroblastes prolifèrent très rapidement et subissent des SHM médiées par AID au niveau de leurs gènes variables afin de générer des cellules B qui expriment des immunoglobulines de haute affinité. La sélection des cellules B ayant une haute affinité à l'antigène s'effectue au niveau de la zone claire en présence des FDC et des T_{FH} (*T Helper Follicular*) ainsi que de cellules T régulatrices (TFR). Plusieurs aller-retours entre les deux zones claire et sombre peuvent être effectués afin de sélectionner les clones les plus affins à l'antigène. L'activation du BCR par un antigène étranger exprimé à la surface des FDC permet la sélection des clones les plus affins. Les BCR faiblement liés à l'antigène meurent par apoptose. Il a été aussi montré que les cellules T_{FH} participent à cette sélection. Les cellules B du GC capturent l'antigène via le BCR et le présentent en association avec le complexe majeur d'histocompatibilité (CMH) aux cellules T_{FH} . Une affinité élevée du BCR est directement

associée à une plus grande capture de l'antigène, et une présentation plus efficace du complexe peptide-CMH aux cellules T_{FH} transmet des signaux de survie à ces cellules. Les BCR ayant une faible affinité pour l'antigène ne permettent qu'une faible interaction avec les cellules FDC et les cellules T_{FH} et seront donc plutôt éliminées ensuite par apoptose (De Silva and Klein, 2015; Kurosaki et al., 2015; Tangye et al., 2013). Au niveau de la zone claire du GC, après avoir sélectionné les meilleurs clones, les cellules B aussi appelées centrocytes subissent une commutation isotypique ou CSR dans le but de changer l'activité effectrice de l'Ig (**Figure 9**) (Liu et al., 1996).

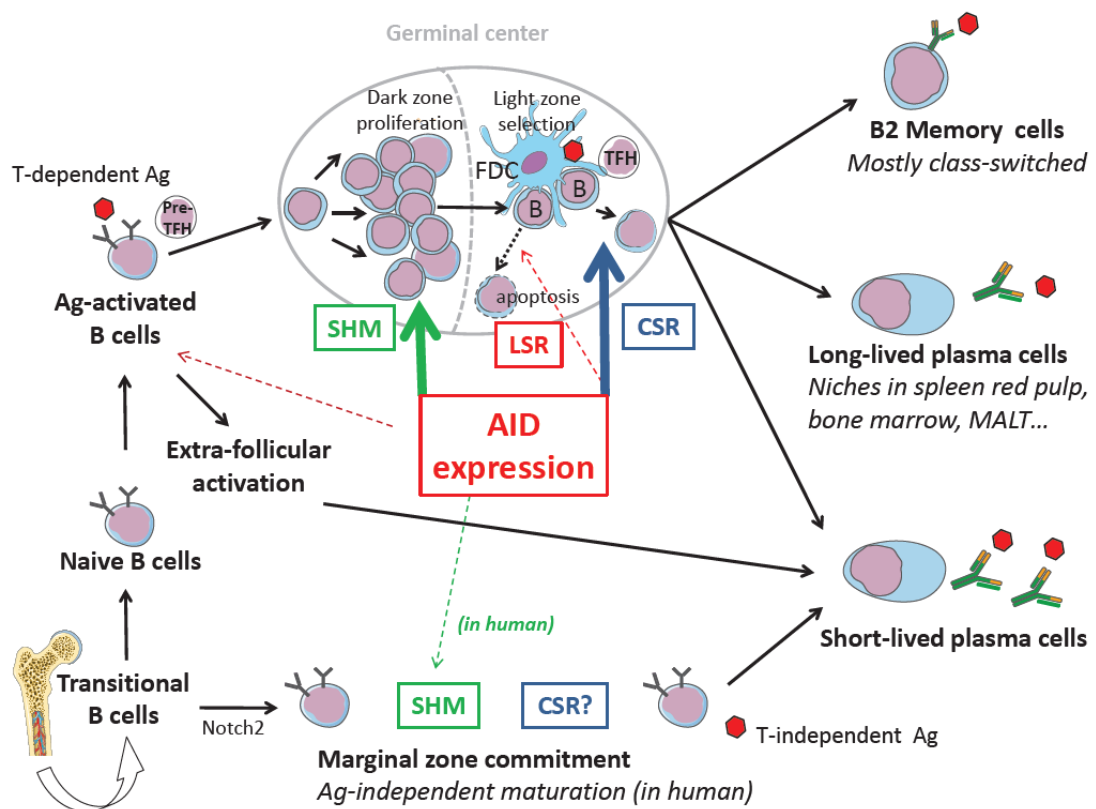


Figure 9: AID et modulation des réactions au sein des centres germinatifs des organes lymphoïdes

Après la recombinaison VDJ et l'expression des IgM, les cellules B transitionnelles peuvent être engagées dans différents compartiments en fonction de la signalisation du BCR (expression Notch2 et faible signalisation tonique BCR permet la différenciation en cellules MZ). Après maturation en cellules IgM + IgD + et la rencontre avec un antigène, les cellules B naïves subissent des interactions apparentées avec les cellules pré- T_{FH} . Les cellules B s'activent ensuite initialement dans les foyers extra-folliculaires et se différencient en plasmocytes à courte durée de vie, ou participent avec les T_{FH} à la formation de GC où leurs BCR seront remodelés par SHM dans les zones sombres. Les cellules B migrent après vers la zone claire pour sélectionner le clone ayant des récepteurs à haute affinité contre l'antigène par l'interaction avec les cellules dendritiques folliculaires (FDC) et les cellules T auxiliaires folliculaires (cellules T_{FH}). Au sein de cette zone les cellules B sous l'action d'AID changent l'isotype de l'Ig exprimée à sa surface par CSR mais peuvent aussi mourir par LSR. La réaction de GC produira à la fois des cellules mémoire et des plasmocytes, certaines ayant une longue durée de vie et survivant après leur migration vers plusieurs niches de protection (pulpe rouge de rate, moelle osseuse, MALT ...)
(Extrait de la revue: Laffleur et al., 2014a).

c. Remaniements secondaires des loci d'Ig induit par AID au sein des centres germinatifs

Suite à une activation antigénique et sous l'action d'AID, la cellule B recombine ses gènes d'immunoglobulines pour diversifier ses fonctions par CSR mais aussi elle augmente l'affinité des Igs par SHM (Muramatsu et al., 2000; Revy et al., 2000). Récemment, notre laboratoire a montré que l'enzyme AID pourrait aussi attaquer la région 3'RR rendant ainsi cette région accessible à des recombinaisons, selon un processus similaire à celui observé pendant la CSR sauf que ces recombinaisons larges entraînent alors la perte du BCR et donc la mort des cellules B par « recombinaison suicide du locus » (LSR) (Péron et al., 2012a).

L'enzyme AID a été découvert en 1999, comme une protéine sélectivement exprimée dans les cellules B activées dans les centres germinatifs (Muramatsu et al., 1999). Il est codé par le gène AICDA localisé sur le chromosome 12 chez l'Homme et sur le chromosome 6 chez la souris (Xu et al., 2007).

AID appartient à la famille des cytidines désaminases APOBEC (*Apolipoprotein B Editing Catalytic Polypeptide*), elle est capable de catalyser la désamination d'une cytosine (dC) en uracile (dU) sur les substrats d'ADN simple brin mais aussi au niveau des hybrides ADN/ARN (Rebhandl et al., 2015). Les points chauds ou « hotspots » de l'AID sont représentés par les motifs RGYW / WRCY (où R = dA / dG, Y = dC / dT et W = dA / dT) (Rogozin and Diaz, 2004). Ces motifs sont ciblés au niveau des gènes variables durant la SHM. Les motifs WRCY sont hautement enrichis dans les régions S, principalement sous la forme d'AGCT (Han et al., 2011). En pathologie, les mutations du gène AICDA entraînent une altération du mécanisme de CSR et de SHM, avec un syndrome d'Hyper-IgM (*HIGM, Hyper-IgM Syndrome*), montrant un rôle majeur d'AID dans la maturation des anticorps (Revy et al., 2000).

La majorité (plus de 90%) de l'AID est localisée dans le cytoplasme mais elle exerce bien sûr sa fonction de désamination dans le noyau des cellules (Vaidyanathan et al., 2014). Le maintien de cette protéine dans le cytoplasme est assuré par la présence d'un site contenant un signal d'exportation nucléaire (*NES, Nuclear Export Signal*) dans la partie C-terminale. L'entrée nucléaire est due à la présence d'un signal de localisation nucléaire (*NLS : Nuclear Localization Signal*) situé dans la partie N-terminale de la protéine (**Figure 10**) (Patenaude and Di Noia, 2010).

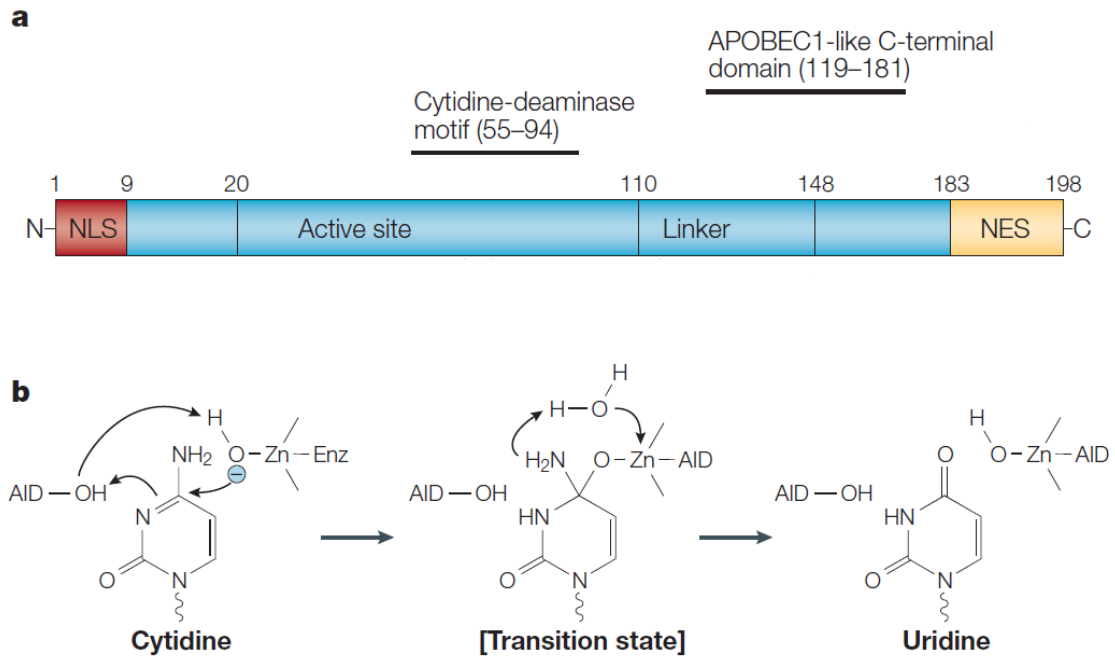


Figure 10: Structure de l'AID et désamination de l'ADN

Structure primaire d'AID (198 acides aminés), représentant le domaine de localisation nucléaire (NLS), le domaine de l'exportation nucléaire (NES), un domaine APOBEC1 et un motif cytidine désaminase (a). AID catalyse la désamination d'une cytidine le transformant en Uracil par attaque nucléophile par un ion Zn^{2+} du groupement amine (b) (Extrait de la revue: Chaudhuri and Alt, 2004).

i. Hypermutation somatique ou SHM

L'hypermutation somatique consiste à introduire des mutations ponctuelles dans la région variable des gènes d'immunoglobuline à un taux de $\sim 10^{-3}$ mutations par paire de bases et par division cellulaire ; ce taux est 10^6 fois plus élevé que le taux de mutation spontanée observé dans les cellules somatiques (Weigert et al., 1970; Yaari et al., 2015). La distribution de fréquence de SHM commence à 100-200 pb en aval du site d'initiation de la transcription du côté 5' du VDJ et s'étend jusqu'à 1.5-2 kb vers le côté 3' (Figure 11) (Peled et al., 2008).

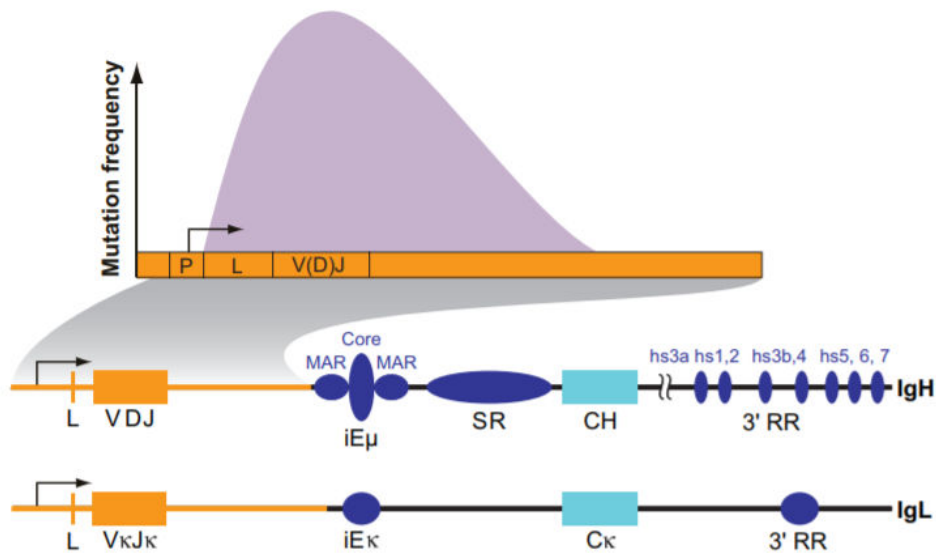


Figure 11: Hypermutation somatique

La fréquence de mutations tout au long des loci IgH et Igk est représentée en violet. Les flèches indiquent le site de début de la transcription. Contrairement à la région constante (en bleu clair) la région V (en orange) montre une fréquence élevée de mutation (SHM). La SHM est nettement délimité par le promoteur (P) situé à l'extrémité 5' et commence à environ 100 à 200 pb après le site de l'initiation de la transcription. La fréquence de mutation est maximale sur l'exon codant V(D)J et décroît de façon exponentielle à l'extrémité 3' à 1,5 - 2 kb en aval du site d'initiation de la transcription.

Leader (L), région V [V(D)J], enhancers introniques (iE μ et iE κ), région C (CH ou C κ), région switch (SR), région 3' régulatrice (3'RR) (*Adaptée de la revue: Peled et al., 2008*)

Après désamination par AID et le remplacement du dC par un dU, le mésappariement G : U peut être réparé par trois « voies » différentes, dont on sait aujourd'hui que les deux dernières sont contemporaines et éventuellement intriquées (Girelli Zubani et al., 2017) :

1) Dans le premier cas, l'uracile peut être ignoré par la machinerie de réplication et elle sera considérée comme une thymidine permettant la transition dC>dT sur un brin et dG>dA sur le brin opposé. La transition est une mutation qui remplace une pyrimidine par une autre pyrimidine ou une purine par une autre purine alors que la transversion est une mutation qui remplace une pyrimidine C ou T par une purine G ou A et vice-versa.

2) La présence anormale d'uracile dans l'ADN peut être traitée par une glycosylase (principalement l'enzyme Uracile ADN glycosylase (UNG), ou encore les enzymes thymidine-DNA-glycosylase ou SMUG1) (Girelli Zubani et al., 2017) de la voie BER (*Base-Excision Repair*). L'uracile excisé crée ainsi un site abasique qui sera ensuite rempli par une des quatre bases (A, T, C, G) pendant la réplication. Cette voie aboutissant ainsi à des mutations de type transition ou transversion au niveau de la paire G-C ciblée par AID.

3) La reconnaissance du mésappariement U : G par l'hétérodimère Msh2 et Msh6 (*Mut S homolog 2 and 6*) de la voie MMR (*MisMatch Repair*) recrute les enzymes MLH1 (*MutL Homologue*) et PMS2 (*PostMeiotic Segregation 2*) ainsi que l'exonucléase EXO1 (*DNA Exonuclease 1*) pour créer une cassure simple brin puis exciser un patch d'ADN incluant la lésion. Par la suite des polymérase de basse fidélité (telle Pol η) sont recrutées pour combler ce patch et éventuellement y créer des mutations à distance de la paire G-C ciblée par AID, tout particulièrement au niveau de paires A : T proches du mésappariement U : G excisé (**Figure 12**) (Bahjat and Guikema, 2017; Odegard and Schatz, 2006; Peled et al., 2008; Rada et al., 2002; Reynaud et al., 2009).

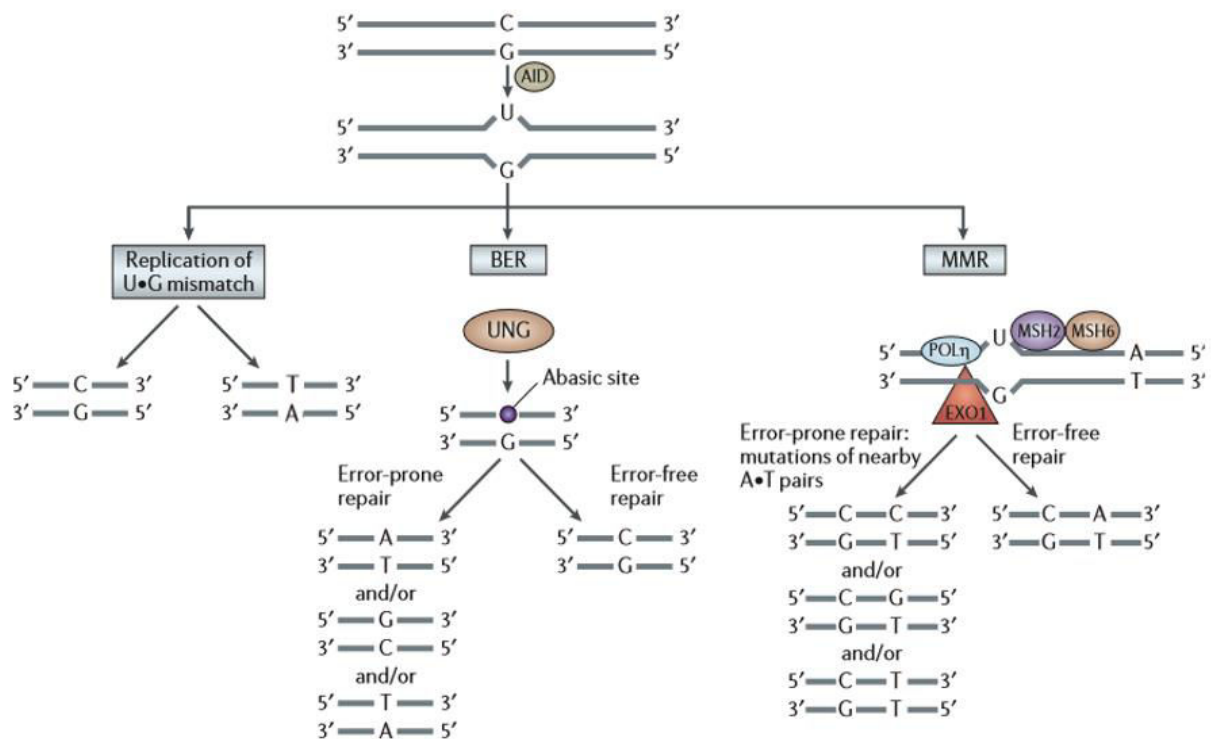


Figure 12: Réparation des lésions induites par AID

L'hypermutation somatique (SHM) s'initie par la désamination de cytidine (C) par AID créant un mésappariement U : G qui sera réparé par trois voies. La machine de réplication peut interpréter le U comme s'il s'agissait d'une désoxythymidine (T). Une des cellules filles va acquérir une mutation de transition C=>T. Pour qu'une mutation soit créée sur le site de désamination, la réparation de l'ADN doit être « perturbée » pour générer des erreurs. UNG peut enlever l'uracile, laissant derrière lui un site abasique. Pendant la réparation par la voie BER, la réplication au niveau du site abasique crée par UNG génère des mutations de transition et de transversion. La réparation du mésappariement par la voie MMR crée des mutations au niveau des paires de base A : T près de la lésion U : G.

UNG, Uracil DNA Glycosylase ; EXO1, exonuclease 1 ; MSH, homologue of Escherichia coli MutS ; POL η , polymérase (DNA directed)- η) (Extraite de la revue: Odegard and Schatz, 2006).

ii. Commutation de classe ou CSR

La CSR permet de substituer la classe initiale d'Ig exprimée par les cellules naïves (IgM ou IgD) par une autre classe qui diffère uniquement par la région constante de la chaîne lourde. Elle s'effectue par recombinaison entre deux régions switch situées en amont de chaque gène constant à l'exception de C δ (Bhattacharya et al., 2010).

La CSR peut être divisée en trois étapes : 1) le choix de la région switch S en aval en tant que cible qui devrait appairer avec la région S μ en amont ; 2) la reconnaissance et le clivage de l'ADN ciblé par AID et 3) la réparation de l'ADN cassé.

L'induction de la CSR nécessite à la fois des stimuli primaires et secondaires. Les stimuli primaires, qu'ils soient T-dépendants (impliquant l'engagement de CD40 exprimé par les lymphocytes B avec le CD154 (nommé aussi CD40L : ligand de CD40) exprimé par les cellules T_{FH}) ou indépendants des lymphocytes T (impliquant les récepteurs TACI (*Transmembrane Activator and CAML Interactor*)) et les récepteurs TLR), activent l'expression d'AID.

Les stimuli secondaires ne peuvent pas induire l'expression d'AID mais ils sont nécessaires pour diriger la commutation de classe vers IgG, IgA ou IgE par une sélection de la région S acceptrice et l'induction de la transcription germinale I_H-S-C_H par une activation de certains facteurs de transcription spécifique selon le type de la stimulation (par exemple IL4, TGF β ...). Cette transcription médiée par l'ARN polymérase II débute au niveau du promoteur de la région I (I_H) du locus IgH et progresse à travers l'exon I_H, la région S et l'exon C_H permettant d'exposer l'ADN simple brin dans les régions S (Haddad et al., 2011; Pone et al., 2012; Xu et al., 2012) créant ainsi une structure appelé boucle R (ou R loop). Cette boucle est une structure dans laquelle le transcrit germinal s'hybride avec le brin d'ADN matrice laissant le brin opposé non apparié sous forme d'ADN simple brin afin d'améliorer l'activité d'AID en lui fournissant son substrat d'ADN simple brin (Gan et al., 2011; Pavri, 2017). Sur le brin opposé non matrice riche en G se forme une structure appelée G-quadruplexe (ou structure G4) (qui peut aussi se former sur l'ARN du transcrit germinal hybridé avec le brin d'ADN matrice). Cette structure G4 est une proie idéale pour AID au niveau de l'ADN ainsi que des transcrits des régions S et elle favorise ainsi la fixation d'AID à la région switch afin d'y exercer sa fonction de désamination mais elle participe aussi à la formation des R-loops grâce à l'intervention d'hélicases comme Ddx1 (Duquette et al., 2005; Ribeiro de Almeida et al., 2018; Zhang et al., 2014; Zheng et al., 2015). La structure « R-loop » est représentée dans la **Figure 13**.

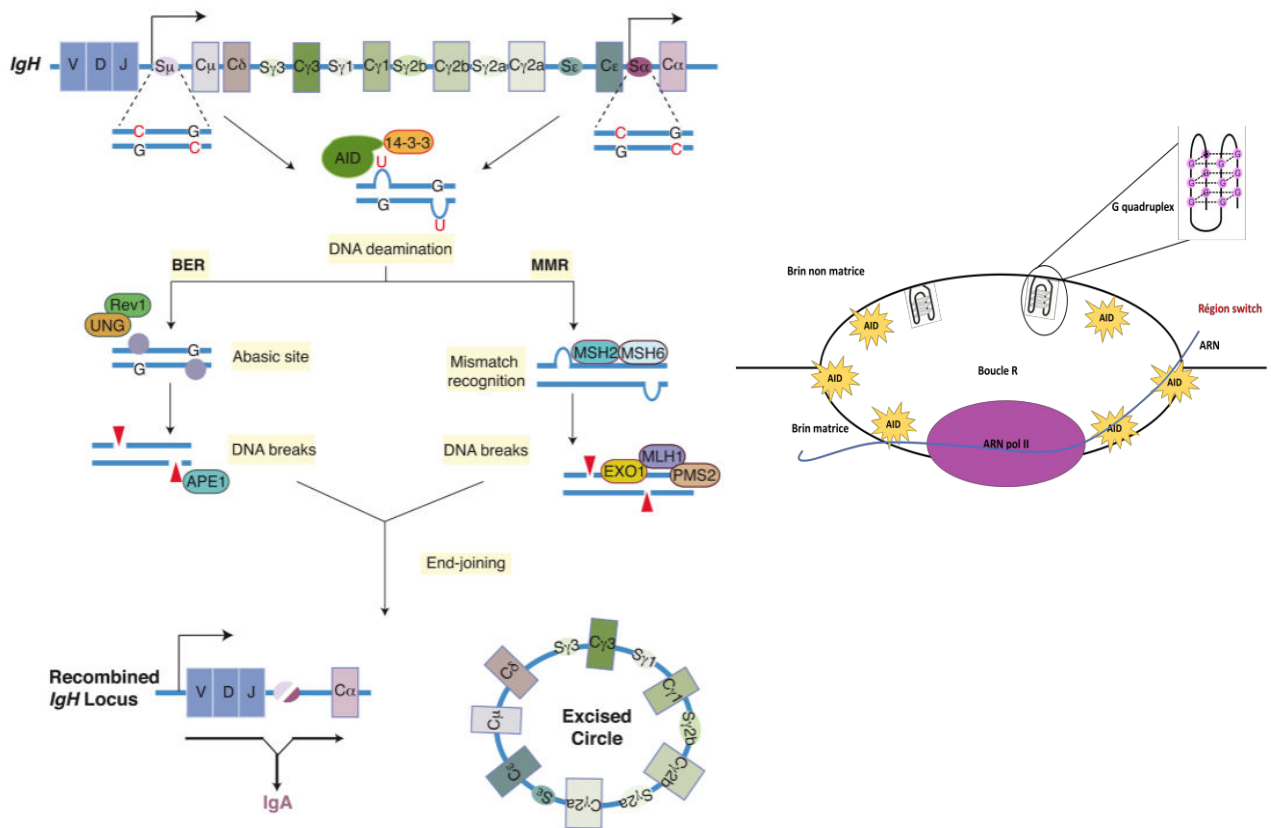


Figure 13: Représentation schématique de la CSR et de la structure de la boucle R

La figure décrit la CSR entre S_{μ} et S_{α} du locus IgH murin. La désamination induite par AID convertit les cytosines en uraciles au niveau des régions switch S. Le mésappariement U : G est converti en cassures double brin d'ADN soit par la voie de réparation par excision de base (BER), soit par la voie de réparation de mésappariement (MMR). Dans la voie BER, UNG (ou d'autres glycosylases) excise la base uracile et génère un site abasique reconnu et clivé par l'endonucléase apurinique / apyrimidinique 1 (APE1). Au cours de la voie de réparation MMR, l'appariement dU : dG est reconnu par le complexe MSH2/ MSH6, qui recrutent le complexe EXO1 / MLH1 / PMS2) pour exciser une courte séquence d'ADN comprenant le mésappariement dU : dG. Après la synapse ADN, la réparation des cassures doubles brins qui joint les deux régions switch S-S est faite par deux voies différentes : la voie de réparation homologue (C-NHEJ) ou par la voie de réparation alternative (A-EJ). Rev1 et 14-3-3 sont des protéines nécessaires à l'assemblage des complexes protéiques participant à la CSR. A droite est représentée la structure de la boucle R qui se forment au niveau des régions switch nécessaire pour la CSR avec la transcription germinale (brin matrice) et la structure en G4 (ou G quadruplex) sur le brin non matrice. La liaison de quatre nucléotides G forme ce qu'on appelle un G-quartet. L'entassement de trois à quatre G-quartets forme le G-quadruplexe facilitant la fixation d'AID (*Adaptée de: Nicolas et al., 2018; Rhodes and Lipps, 2015; Ribeiro de Almeida et al., 2018*).

D'autre part, les modifications post-traductionnelles des histones au niveau de la région *switch* donneuse en amont et la région *switch* acceptrice en aval jouent également un rôle crucial pour initier la transcription germinale. La transcription de I_{μ} - S_{μ} - C_{μ} marqué par la présence des histones activées, tels que H3K4me3, H3K36me3, H2BK5ac, H3K9ac / K14ac, H3K27ac et H4K8ac (associé à la présence d'une chromatine active) est observée dans la région S_{μ} même

dans les cellules B naïves au repos, ce qui suggère que $S\mu$ est dans un état constitutivement ouvert et prêt pour la recombinaison. Les stimuli primaires activent les enzymes qui catalysent des modifications des histones (acétylation ou méthylation) tels que les histones méthyltransférases et acétyltransférases pour déclencher la décondensation de la chromatine dans la région S en aval donnant l'accès à l'ARN pol II (Li et al., 2013; Zan and Casali, 2015).

La désamination de la cytosine par AID crée des lésions au niveau des deux régions *switch* acceptrice et donneuse. Cette étape se déroule de la même façon que la SHM. Le mésappariement U : G est prise en charge par les deux voies BER et MMR aboutissant à la formation des cassures simples brins SSB (*Single Strand Break*) puis doubles brins dans les deux régions *switch* (Hackney et al., 2009; Xu et al., 2012). Ces cassures excisent l'ADN situé entre ces deux régions sous forme d'un ADN circulaire (épisode) alors que les deux régions *switch* sont liées et forment une synapse $S\mu$ - $S\alpha$ permettant la juxtaposition d'un nouvel exon CH d'une autre classe qui sera exprimée à la membrane des lymphocytes B (**Figure 13**) (Matsuoka et al., 1990; von Schwedler et al., 1990).

Les cassures doubles brins générées au niveau de l'ADN sont potentiellement dangereuses et peuvent être la cause de translocations génomiques potentiellement oncogéniques. Pour cette raison, les cassures doivent être rapidement détectées et réparées. Pendant la CSR, deux voies distinctes sont connues pour réparer ces cassures : la voie de réparation non homologue NHEJ et la voie de réparation alternative A-EJ (Kotnis et al., 2009).

- **Réparation des cassures doubles brin par la voie non homologue NHEJ**

Les DSB générées au niveau des régions *switch* activent et recrutent plusieurs facteurs qui jouent un rôle important dans la réparation de ces lésions. Dans un premier temps, le complexe formé par Mre11-Rad50-Nbs1 (MRN), se localise sur les DSB et recrute et phosphoryle la protéine ATM (*Ataxia Telangiectasia Mutated*) dont le rôle principal est de recruter et de phosphoryler des protéines de la réparation de l'ADN telles que l'histone H2AX (*H2A histone family member X*) qui facilite la formation de la synapse entre les deux régions *switch*, la protéine NBS1 (*Nijmegen Breakage Syndrome 1*) et d'autres protéines comme la protéine 53BP1 (*p53 Binding Protein 1*) et la protéine MDC1 (*Mediator of DNA Damage Checkpoint 1*). Plusieurs études ont montré que des mutations dans ces protéines sont accompagnées avec une augmentation des translocations génomiques.

Lors de la réparation par cette voie, les hétérodimères Ku70/80 se lient aux extrémités de l'ADN cassé et recrutent la protéine kinase ADN-PKs. L'hétérodimère Ku70/86 et l'ADN-PKs sont

impliqués dans la formation de la synapse S-S. L'ADN-PKs active aussi sa propre sous-unité catalytique (ADN-PKcs) et recrute l'endonucléase Artemis, au niveau de la DSB pour enlever les séquences d'ADN simple brin en excès. Enfin le complexe XRCC4 et l'ADN ligase IV catalyse la ligation des deux régions *switch* (**Figure 14 à gauche**) (Kotnis et al., 2009; Li and Xu, 2016).

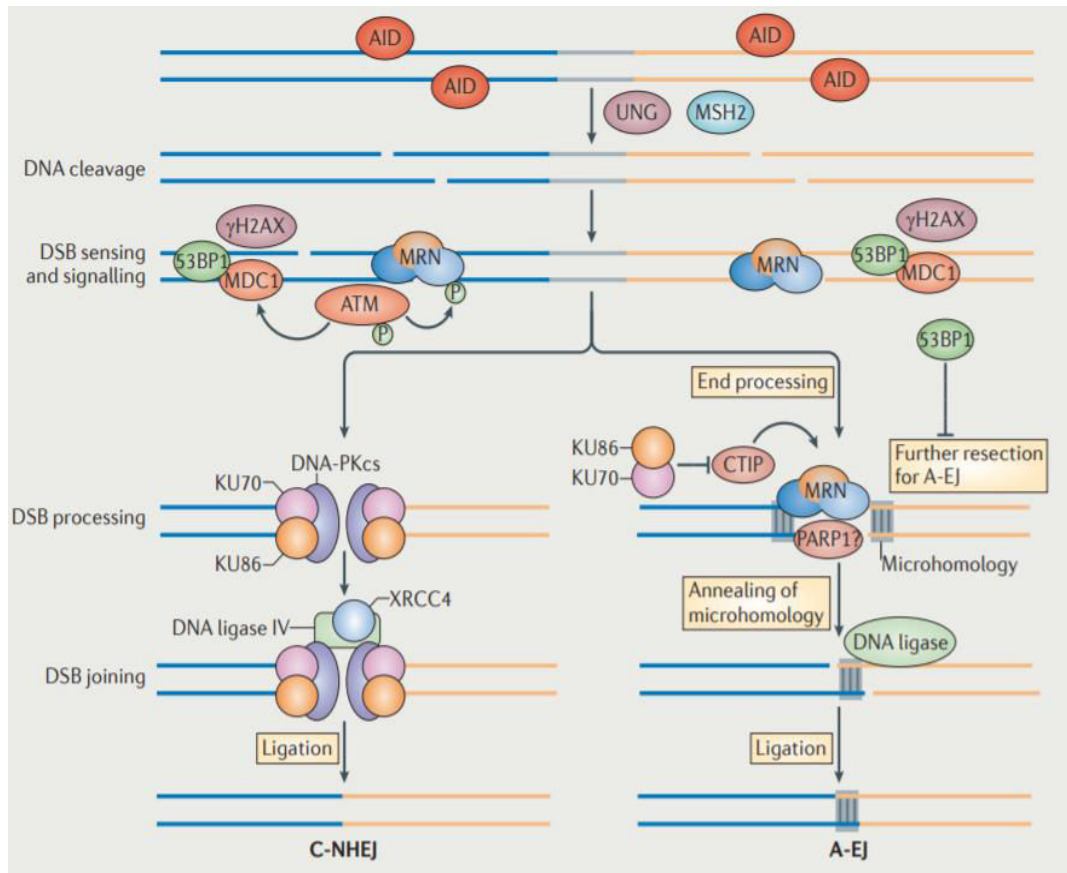


Figure 14: Réparation des cassures par les voies NHEJ et A-EJ

Les cassures d'ADN générées par AID sont résolues par deux voies : La voie NHEJ (à gauche) et la voie A-EJ (à droite).

C-NHEJ: Classical Non-Homologous End Joining; A-EJ: Alternative End joining (Extrait de Xu et al., 2012).

- **Réparation des cassures doubles brins par la voie alternative A-EJ**

Des études faites sur des souris SCID (*Severe Combined Immunodeficiency*) ayant l'ADN-PK non fonctionnelle ne montrent pas une forte diminution de la CSR en comparaison avec des souris WT suggérant qu'il y a une voie supplémentaire qui pourrait compenser l'activité de l'ADN-PKs, il s'agit de la voie alternative (Bosma et al., 2002).

La voie A-EJ fait intervenir également le complexe MRN et la protéine CTIP (*CtBP-Interacting Protein*) pour résoudre les cassures d'ADN. Elle utilise comme matrice une séquence d'ADN identique à la séquence endommagée pour la liaison des deux régions *switch* (jonctions avec 4

à 20 pb de microhomologie) alors que la voie NHEJ n'utilise pas des microhomologies entre les deux régions *switch* pour réparer les cassures d'ADN où les jonctions S μ -S x créés sont obtenues par des liaisons franches (jonction directe) mais on trouve aussi des jonctions avec des courtes microhomologies (<4pb).

L'analyse de la structure des jonctions nous permet donc de déduire la voie de réparation la plus probablement impliquée (Stavnezer et al., 2010). L'étude des jonctions de CSR chez des patients ayant une déficience pour la Ligase IV montre une absence de jonctions directes et la présence de longues séquences de microhomologie (4 à 20 pb) mais aussi une augmentation des jonctions avec insertion grâce à l'utilisation de la voie A-EJ (**Figure 14 à droite**) (Pan-Hammarström et al., 2005).

- **Recombinaison *trans* ou *trans-Switch* : cas particulier de CSR**

Bien que la CSR se produise majoritairement par une recombinaison intrachromosomique entre deux régions *switch* situées sur le même allèle (*cis*-CSR), Kingzette et ses collègues ont montré de longue date chez le lapin (qui possède des allotypes de région variable) que des recombinaisons en *trans* surviennent entre les deux allèles (*trans*-CSR) induisant la production d'IgA dont les régions VH et C α associées sur la chaîne lourde sont pourtant codées par des gènes initialement situés en *trans*. Ce type de recombinaison sur le locus IgH se produit chez le lapin avec une fréquence élevée de 7% (Kingzette et al., 1998). Ces recombinaisons inter-alléliques sont en fait un processus général et ont été aussi mises en évidence chez la souris ; elles se produisent avec une fréquence de 7% à 17% pour le *switch* vers IgA et IgG3 (Reynaud et al., 2005) et jusqu'à 15-20% pour le *switch* vers IgA1 chez l'Homme (Laffleur et al., 2014b).

Lors de la CSR, les promoteurs I subissent des interactions à longue distance avec la région 3'RR afin de répondre à certaines stimulations permettant d'activer la transcription et l'accessibilité des régions *switch*. Ces modifications mais aussi d'autres pré-requis de la CSR (telle l'interaction avec des facteurs de transcription, les modifications chromatinienne dans les régions *switch*...) se produisent à la fois sur l'allèle « productif » et l'allèle « non productif » du locus IgH, permettant l'apparition de cassures d'ADN double brin sur les deux allèles. Ceci explique à la fois que le *switch* survienne en règle parallèle sur les 2 allèles IgH que soient générés des sites potentiels pour des recombinaisons *trans*-chromosomiques (Borzillo et al., 1987; Delpy et al., 2003; Dougier et al., 2006; Khamlichi et al., 2000; Rothman et al., 1990).

4. Stade de différenciation finale : Sécrétion des anticorps et mémoire immunitaire

La dernière étape de différenciation et de la lymphogenèse B est marquée par la génération des plasmocytes sécréteurs d'anticorps qui sont les acteurs de l'immunité humorale et des cellules B mémoires qui assurent une réponse plus efficace et plus rapide lors d'une nouvelle rencontre avec le même antigène.

La différenciation des cellules B en plasmocytes dépend de l'environnement, la nature de l'antigène et de la population des cellules B dont elles sont issues. Les plasmocytes à courte durée de vie sont notamment issus des cellules B de la zone marginale et des cellules B1 indépendamment de l'aide des cellules T, mais aussi de l'activation B T-dépendante dans les régions extra-folliculaires pour les cellules portant un BCR de faible affinité pour l'antigène (Oracki et al., 2010; Sacquin et al., 2017). Ces cellules assurent une réponse immédiate contre l'antigène. Les plasmocytes à longue durée de vie sont générés à partir des cellules B folliculaires reconnaissant l'antigène à haute affinité et activées de façon TFH-dépendante. Ces cellules peuvent migrer ensuite vers diverses niches (principalement la moelle osseuse en condition physiologique, mais aussi la rate ou d'autres sites lymphoïdes activés au cours de maladies auto-immunes) et y assurer une protection pouvant durer pour des années (Nutt et al., 2015; Shapiro-Shelef and Calame, 2005; Thai et al., 2018). Cette différenciation plasmocytaire est associée à certaines modifications du profil d'expression génétique qui permettent la répression des facteurs B spécifiques tels que Pax5 et l'expression des facteurs plasmocytaires spécifiques (IRF4 (*Interferon Regulatory Factor 4*), XBP1 (*X-box Binding Protein 1*), BLIMP1 (*B Lymphocyte-Induced Maturation Protein 1*)) afin d'activer le programme de la sécrétion d'anticorps (Kassambara et al., 2017; Tellier et al., 2016).

Les cellules mémoires B résident dans les organes lymphoïdes secondaires, et qui peuvent être rapidement mobilisées lors d'une nouvelle rencontre antigénique. Après réactivation, ces cellules se différencient majoritairement en plasmocytes de longue durée et sécrètent des Ac qui peuvent être détectés parfois pendant des années. Il a été suggéré que les cellules B mémoires soient maintenues par la présence prolongée de l'antigène et qu'elles pourraient être stimulées de façon continue pour donner lieu à des plasmocytes de courte durée contribuant à maintenir le niveau de anticorps dans le sérum. D'autres études ont montré que les cellules B mémoire pourraient survivre dans un état quiescent en l'absence d'antigène (Kurosaki et al., 2015; Reynaud et al., 2012).

IV. Activateurs transcriptionnels du locus IgH

Les activateurs (*enhancer*) et les super-activateurs (*superenhancer*) transcriptionnels sont les principaux régulateurs du destin cellulaire qui agissent sur des longues distances sur des gènes adjacents. Le locus de la chaîne lourde des immunoglobulines IgH portent deux *superenhancers* situés sur les deux extrémités 5' et 3' des gènes de la région constante. Le *superenhancer* 5'E μ qui favorise la recombinaison VDJ pendant les premières étapes de l'ontogenèse des cellules B et le *superenhancer* 3'RR (*3'regulatory region*) ou la région régulatrice 3'RR qui joue un rôle essentiel durant les phases de différenciation tardive (Saintamand et al., 2017).

D'autres activateurs d'importance moindre ont été identifiés chez la souris à différentes positions du locus, parmi les segments V et au voisinage du segment DQ52.

1. L'activateur intronique E μ

C'est le premier *enhancer* découvert en 1983 au sein du locus IgH (Banerji et al., 1983; Gillies et al., 1983). Il est localisé en amont des gènes constants dans l'intron entre le dernier segment J_H (JH4 pour la souris et JH6 chez l'Homme) et le gène constant C μ . Chez la souris, il est organisé sous forme d'un élément central de 220 pb appelé core E μ ou cE μ flanqué de deux régions d'attachements à la matrice nucléaire ou Matrix Attachment Regions (MARs) de 310 à 350 pb. Le cE μ est formé de plusieurs sites de liaison de facteurs de transcription. De 5' en 3' se situe un site C/EBP (*Ccaat-Enhancer-Binding Proteins*), μ E1, μ E5, μ E2, μ A, μ E3, μ E4 et une séquence octamérique. Cette séquence octamérique permet la fixation de certains facteurs B spécifiques Oct1 et Oct2 (Ernst and Smale, 1995)

La délétion complète d'E μ (core et matrice) chez la souris bloque partiellement la transition du stade de différenciation pro-B vers le stade pré-B avec une accumulation de la proportion et du nombre des précurseurs pré-pro-B et des cellules pro-B. La délétion d'E μ diminue mais n'abolit pas la CSR, alors qu'elle affecte drastiquement l'assemblage des segments VDJ (Marquet et al., 2014).

2. La région régulatrice 3'RR

a. Structure de la 3'RR

La région régulatrice 3' du locus IgH ou 3'RR s'étend sur une région de \sim 30 kb localisée en aval du gène C α et joue un rôle crucial pendant la CSR et la synthèse d'Ig mais aussi pendant l'hypermutation somatique SHM qui touche les gènes de la région variable. Le locus IgH de la souris possède une seule copie de la région 3'RR alors que le locus IgH humain possède deux

copies nommées 3'RR1 et 3'RR2 localisées en aval de Ca1 et Ca2 respectivement. L'existence de deux copies de 3'RR est due à la duplication du locus IgH, ces régions possèdent une séquence très similaire mais avec une large inversion d'un segment central et l'existence d'allèles de la 3'RR1 liés à des variations d'orientation et de séquence de l'activateur hs1,2 (Frezza et al., 2012; Pinaud et al., 1997). Chaque région 3'RR humaine contient 3 activateurs qui constituent des sites hypersensibles à la DNase I (hs1,2, hs3 et hs4). La seule région 3'RR chez la souris contient quatre sites hypersensibles à la DNase I (hs1,2, hs3a, hs3b et hs4). Trois autres sites ont été également identifiés en aval de la 3'RR murine et nommés hs5, hs6 et hs7 (**Figure 15**) (Birshtein, 2014; D'Addabbo et al., 2011; Pinaud et al., 2011).

Une analyse de séquence sur 23 espèces différentes de mammifères montre que tous possèdent au moins une copie de la 3'RR mais seulement 8 espèces présentent les 3 activateurs hs1,2, hs3 et hs4 ensemble (Chat, Chien, Panda, Gorille, Lapin, Souris, Rat noir et Orang-Outan). Les dauphins ne semblent pas posséder hs4, les chevaux ne semblent pas posséder hs1-2 ni hs4. Des homologues des isolateurs hs5, hs6 et hs7 présents chez la souris en aval de la 3'RR n'ont été retrouvés que chez le rat (D'Addabbo et al., 2011). Il est cependant à noter que ces comparaisons sont généralement faites *in silico* et qu'elles peuvent donc tout à fait négliger des éléments non codants dont la séquence aurait fortement divergé tout en conservant sa fonction, d'autant qu'un certain nombre de facteurs de transcription fixant la séquence des *enhancers* de la 3'RR ont des séquences assez dégénérées (cas de Pax5 par exemple).

Chez la souris, les trois premiers éléments de cette région hs3a, hs1,2 et hs3b et leurs régions flanquantes forment une structure dite palindromique qui s'étend sur 25 kb et dont hs1-2 occupe le centre. hs1-2 est donc entouré de part et d'autre par des séquences répétées et inversées (*IR* : *Inverted Repeats*) de 10 à 12 kb (certaines sont en tandem). Les éléments hs3a et hs3b occupent les extrémités de ce palindrome, ils partagent 97% de similitude et ils ont des orientations inverses (Chauveau and Cogné, 1996). L'élément hs4 situé à 4 kb en aval du palindrome et jouerait son rôle d'*enhancer* selon un mode un peu différent des autres *enhancers* 3' (Michaelson et al., 1995). L'ensemble de la région régulatrice s'étendant entre ces 4 *enhancers* joue un rôle global dans la fonction et constitue fonctionnellement un *superenhancer*, dont l'architecture globale (et sa nature palindromique) contribuent à l'efficacité des recombinaisons (Le Noir et al., 2017a; Saintamand et al., 2016). Cette région *superenhancer* représente environ ~28 kb et est suivie par une autre région de ~12 kb formée par hs5, hs6 et hs7 riches en sites de fixation pour CTCF (*CCCTC-binding factor*) et Pax5, qui joue le rôle d'isolateur (mais dont la délétion par knock-out n'a révélé qu'un effet très modeste sur l'usage des V, sans impact sur

les recombinaisons AID-dépendantes) (Birshtein, 2014; Pinaud et al., 1997, 2011; Volpi et al., 2012).

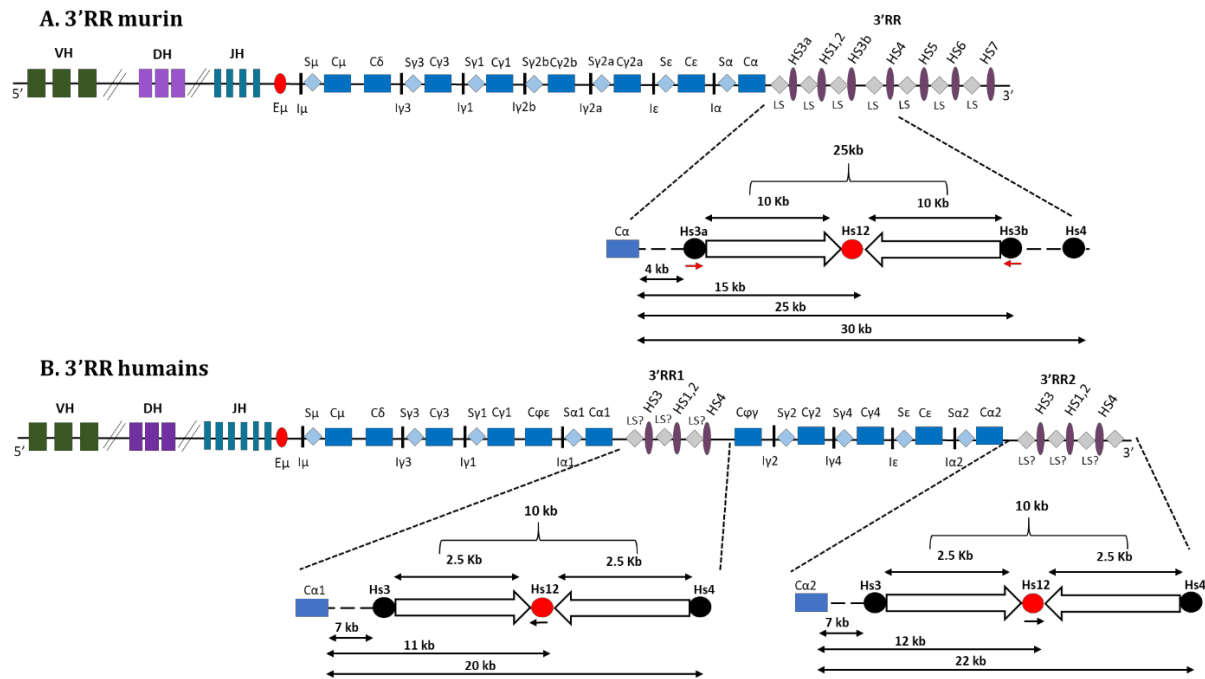


Figure 15: Structure de la 3'RR chez l'Homme et chez la souris

A. La souris possède une seule région 3'RR possédant 4 sites hypersensibles à la DNase I flanqués par des régions LS notés de 5' vers 3' : hs3a, hs1-2, hs3b et hs4 à 4 kb, 15 kb, 25 kb et 30 kb de Cα respectivement. Le palindrome est formé comme chez l'Homme de séquences répétées et inversées de 10 kb de chaque côté d'hs1-2. hs3a et hs3b situées aux extrémités de ce palindrome ont une orientation inversée. Ces activateurs hs sont flanqués par des régions Like-switch LS représentées par des losanges gris.

B. L'Homme possède deux 3'RR nommées 3'RR1 et 3'RR2. Chaque région possède trois sites hypersensibles à la DNase I noté de la côté 5' vers 3' : hs3, hs1-2 et hs4 à 7kb, 11 kb et 20 kb de Cα1 respectivement et à 7 kb, 2 kb et 22 kb de Cα2 respectivement. Le palindrome est formé de séquences répétées et inversées de 2.5 kb qui sont représentées par des flèches ayant une orientation inversée. hs3 et hs4 occupent les extrémités de ce palindrome avec hs1-2 au centre. La séquence d'hs1-2 est inversée dans la 3'RR1 par rapport à la 3'RR2.

Les deux 3'RR chez l'Homme mesurent chacune environ 20 kb, et incluent chacune (comme chez la souris même si les séquences des régions intervenantes, non codantes, ont totalement divergé) une structure palindromique d'environ 9-10kb, donc taille, un peu plus courte que le palindrome observé chez la souris. Cette structure de 10 kb dont hs1-2 occupe aussi le centre est flanquée du côté 5' par hs3 et du côté 3' par hs4. La similarité entre les deux versants du palindrome est de l'ordre de 94%. La présence de hs1-2 au milieu du palindrome avec deux

éléments de segments répétés pourrait faire postuler que le *superenhancer* adopte éventuellement une conformation tri-dimensionnelle en épingle à cheveux de cette région, même s'il reste à comprendre comment cet arrangement peut optimiser l'effet du *superenhancer*. Au centre, hs1-2 peut avoir ou non (selon un allélisme de la 3'RR1) une orientation opposée à celle de l'élément hs1-2 de la 3'RR2 (D'Addabbo et al., 2011; Pinaud et al., 1997, 2011). Après la description initiale de ces allèles de la 3'RR1 par notre équipe et de son impact sur le niveau de production d'IgA1 au cours des néphropathies à IgA (Aupetit et al., 2000), les travaux du groupe de D. Frezza ont connecté cet allélisme à la prévalence de diverses autres manifestations dys-immunitaires (Cianci et al., 2008; Frezza et al., 2004, 2012).

Outre la fonction d'*enhancer*, notre laboratoire a décrit l'existence des séquences répétées ressemblant à celles trouvées dans la région Sy1 chez l'Homme et chez la souris (Chauveau and Cogné, 1996; Péron et al., 2012a). Ces régions, nommées « Like-Switch » (LS) peuvent constituer des cibles potentielles pour le mécanisme de la recombinaison suicide du locus IgH (LSR) (Péron et al., 2012a).

b. Activité des *enhancers* de la 3'RR lors de l'ontogénie B et les remaniements qui touchent le locus IgH

La région 3'RR ne semble pas nécessaire à la recombinaison VDJ et à l'établissement du répertoire pré-immun (Rouaud et al., 2012). Elle démontre plutôt même un effet *silencer* de la transcription IgH dans les progéniteurs B (Braikia et al., 2015). La région 3'RR est le régulateur principal du locus IgH qui contrôle la commutation de classe CSR, l'hypermutation somatique SHM mais aussi l'expression du locus IgH durant le stade plasmocytaire (Le Noir et al., 2017a). L'activateur hs4 mais aussi les activateurs hs5, hs6 et hs7 deviennent actifs pendant le stade pro-B et pré-B et qui reste actif tout au long du développement B alors que hs1-2, hs3a et hs3b n'apparaissent actifs que pendant les stades les plus avancés surtout dans les cellules plasmocytaires. Cette activation est associée avec une chromatine active (acétylation de l'histone H3 et H4, déméthylation de la lysine 9 d'H3 et déméthylation de l'ADN) créant une conformation ouverte de l'ADN permettant la fixation de certains facteurs de transcription (Garrett et al., 2005; Giambra et al., 2008).

L'implication de la 3'RR dans les remaniements qui se produisent au locus IgH a été étudiée sur différents modèles de souris et de lignées cellulaires. La première délétion de la région 3'RR alors même que sa séquence, sa structure et son rôle régulateur étaient inconnus (et donc sans que le phénotype obtenu ne puisse alors être compris), a été en fait observée il ya plus de 30

ans dans une lignée de myélome murine sécrétrice d'IgA. Cette lignée présentant une délétion spontanée d'une large séquence en aval de C α montrait une diminution de la transcription d'IgA (Gregor and Morrison, 1986). Plus tard, Lieberson et ses collègues ont montré que le remplacement d'hs1-2 par un gène de *néo* dans une lignée cellulaire sécrétrice d'IgG2a et déjà déficiente en E μ affecte la transcription du locus IgH alors que cette même lignée déficiente en E μ continue à sécréter l'IgG2a montrant que cette production d'Ac est 3'RR dépendante (Lieberson et al., 1995a). La fonction de la 3'RR dans les cellules B a été approchée par des expériences d'inactivation dans des cellules ES. Le remplacement ciblé de hs1-2 et d'hs3a par un gène de néomycine a d'abord suggéré son implication majeure dans la transcription germinale de plusieurs gènes de la région constante du locus IgH et dans la CSR (Cogné et al., 1994; Manis et al., 1998). Cependant ce phénotype observé serait dû moins à la délétion ponctuelle d'un unique *enhancer* qu'à l'effet « *Néo* », puisqu'une nouvelle délétion d'un gène de résistance à la néomycine « floxé » restaure la CSR (Manis et al., 1998; Pinaud et al., 2001). En revanche et même sans insertion de cassette neo^R, la délétion combinée des deux activateurs hs3b et hs4 chez la souris altère la transcription germinale accompagnée avec une baisse de la synthèse des Igs sauf pour IgG1 et IgM (Pinaud et al., 2001). Il a été suggéré que cet effet est dû à la délétion d'hs4 et pas à hs3b parce que la délétion d'hs3b qui partage 93% de similarité avec hs3a et qui est absent chez l'Homme portant donc une sorte de knock-out naturel de hs3b n'a pas pour autant de « handicap » en matière de CSR. Pour clarifier ce point, notre laboratoire a développé deux modèles de souris qui possèdent une délétion d'hs4 ou d'hs3a et un troisième modèle qui possèdent une délétion de l'intégralité de la 3'RR afin de savoir si ces activateurs pourraient être impliqués individuellement ou nécessitent une coopération afin de réguler l'expression du locus IgH. Seules les souris ayant une délétion de toute la région 3'RR montrent une diminution de la synthèse de tous les isotypes due à la diminution de la transcription du locus IgH (incluant IgM) mais aussi une altération de la CSR. Ce phénotype est aussi accompagné par une altération du mécanisme de l'hypermutation somatique SHM alors qu'elle n'affecte pas la recombinaison VDJ soulignant que cette région pourrait orchestrer l'activité du locus IgH au cours des derniers stades de la différenciation des cellules B. Les seuls effets observés avec la délétion d'hs4 sont une diminution du nombre des splénocytes et de l'expression des IgM membranaires. La délétion de toute la 3'RR pourrait perturber la boucle qui se forme entre E μ , les régions switch et la 3'RR nécessaire pour la CSR (Bébin et al., 2010; Rouaud et al., 2012, 2013; Vincent-Fabert et al., 2009, 2010). Après que de multiples délétions, insertions ou restructurations de la 3'RR aient été analysées par notre laboratoire, l'image

globale qui se dégage est sans doute que toute altération de la 3'RR affecte les remaniements tardifs du locus IgH (CSR et/ou SHM), mais soit de façon discrète pour des délétions ponctuelles (même quand celles-ci touchent un « core *enhancer* »), ou soit de façon majeure quand il s'agit de délétions larges (plusieurs *enhancers*) ou d'insertions de gènes exogènes (comme neoR) ou encore quand tous les core enhancers sont respectés mais qu'on a « altéré » l'architecture des séquences intervenantes en répétition inversées séparant ces enhancers (Le Noir et al., 2017a; Pinaud et al., 2011; Saintamand et al., 2016).

c. Implication de la 3'RR dans la structure tridimensionnelle du locus IgH

L'expression des gènes est régulée par des interactions entre des éléments régulateurs séparés par des centaines de kilobases. Plusieurs études ont montré que des éléments chromosomiques éloignés dans de nombreux loci génétiques s'associent pour former des boucles chromatiniennes fournissant ainsi un mécanisme d'activation permettant l'expression des gènes. Les régions S ciblées pour la recombinaison sont situées entre les deux activateurs $E\mu$ et 3'RR pouvant être séparées par plus que 150 kb ; cette distance ne s'avère pas un obstacle à la formation des synapses S-S (Feldman et al., 2015; Wuerffel et al., 2007). Dans les cellules B au repos, l'interaction entre $E\mu$ et la 3'RR favorise la formation d'une boucle permettant la juxtaposition de ces deux régions alors que l'activation des lymphocytes B conduit à la juxtaposition des deux régions *switch* dans la même boucle permettant la transcription germinale et la recombinaison (Kenter et al., 2013). La délétion de la 3'RR altère la formation de la boucle induite par l'interaction entre $E\mu$ et la 3'RR montrant que la CSR est dépendante de la structure tridimensionnelle de la chromatine facilitant des interactions intrachromosomiques à longue distance entre des éléments transcriptionnels éloignés (**Figure 16**) (Feldman et al., 2015; Vincent-Fabert et al., 2010). Il faut enfin garder un regard critique sur le rôle de ces boucles, car même si leur existence physique est démontrée elles peuvent fort bien n'être qu'une conséquence des synapses entre régions *switch*, et non un préalable à ces synapses ; la haute fréquence des recombinaisons inter-alléliques observées en physiologie plaide plutôt pour une telle vision (Dougier et al., 2006; Laffleur et al., 2014b; Reynaud et al., 2005). En ce sens, des régions riches en AID actif (notion d'AID factories » ou de « *switch factories* ») pourraient fort bien jouer un rôle fonctionnel et constituer un attracteur formant aussi bien la base de boucles intra-chromosomiques, qu'un lieu pour les recombinaisons alternatives (mais tout aussi légitimes) de *trans-switch* entre allèles IgH. L'existence d'interactions physiques et fonctionnelles entre allèles IgH, favorisées par la 3'RR, va dans ce sens (Le Noir et al., 2017b).

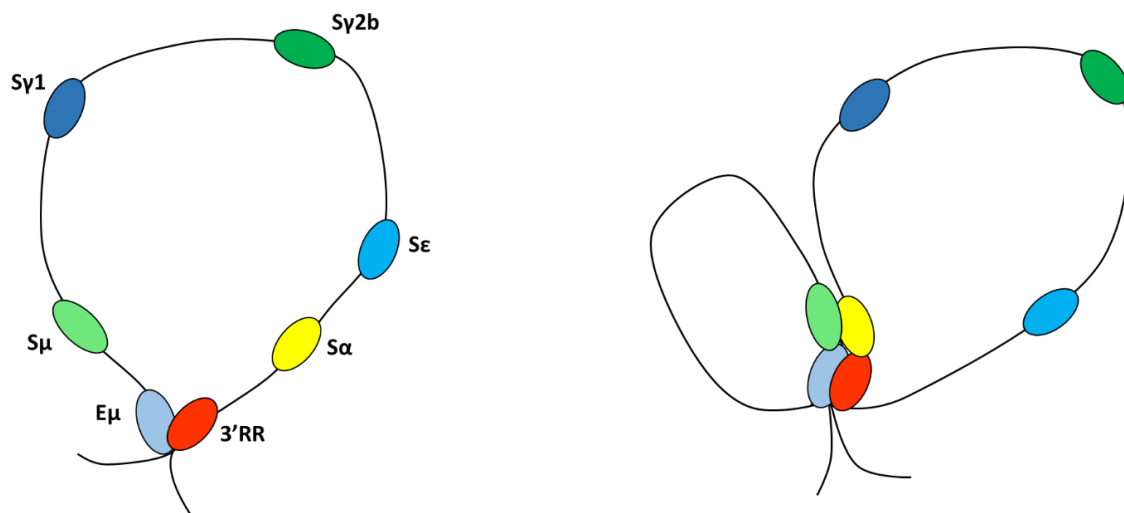


Figure 16: Structure tridimensionnelle du locus IgH facilitant la CSR dans les cellules B matures

Dans les cellules B, une interaction entre $E\mu$ et la $3'RR$ créerait (classiquement et au moins dans le cas de la *cis*-CSR) une structure tridimensionnelle du locus IgH (à droite). Lors de l'activation des cellules B la juxtaposition de $S\mu$ et de $S\alpha$ favorise la formation de la synase S-S.

d. La région $3'RR$: cible pour la recombinaison suicide du locus IgH ou LSR

Un nouveau mécanisme de recombinaison médiée par AID, appelée la recombinaison suicide du locus IgH (LSR) a été mis en évidence en 2012 chez la souris par notre laboratoire. C'est une recombinaison entre la région $S\mu$ et la région $3'RR$, qui délète l'ensemble des gènes constants et aboutit à la mort des cellules B par la perte de l'expression du BCR à la membrane cellulaire (Péron et al., 2012a).

Les activateurs transcriptionnels 3' du locus IgH sont flanqués par des séquences répétitives en tandem qui ressemblent aux régions *switch*, ces régions sont appelées LS (Chauveau and Cogné, 1996). L'analyse des dots plots de la $3'RR$ de la souris a révélé 11 segments de LS de 0,5 à 1 kb dont 60% contiennent les mêmes répétitions qui se trouvent au niveau de $S\gamma 1$ (blocs de répétition de 49 pb). Du côté 5' vers 3', LS1 précède $hs3a$; six segments (LS2 à LS7) sont inclus dans le palindrome flanquant $hs1-2$, deux segments (LS8 et LS9) sont situés après $hs3b$, et deux autres (LS10 et LS11) sont situés en aval de $hs4$. Les dots plots de la $3'RR2$ chez l'Homme montrent la présence de six régions LS : LS1 précède $hs3$, deux segments (LS2 et LS3) précèdent $hs1,2$; LS4 et LS5 précèdent $hs4$ et enfin LS6 est situé en aval d' $hs4$. La présence de ces régions LS au niveau de la $3'RR$ suggère qu'elles peuvent avoir un rôle fonctionnel qui pourrait être une cible potentielle pour AID (**Figure 17**).

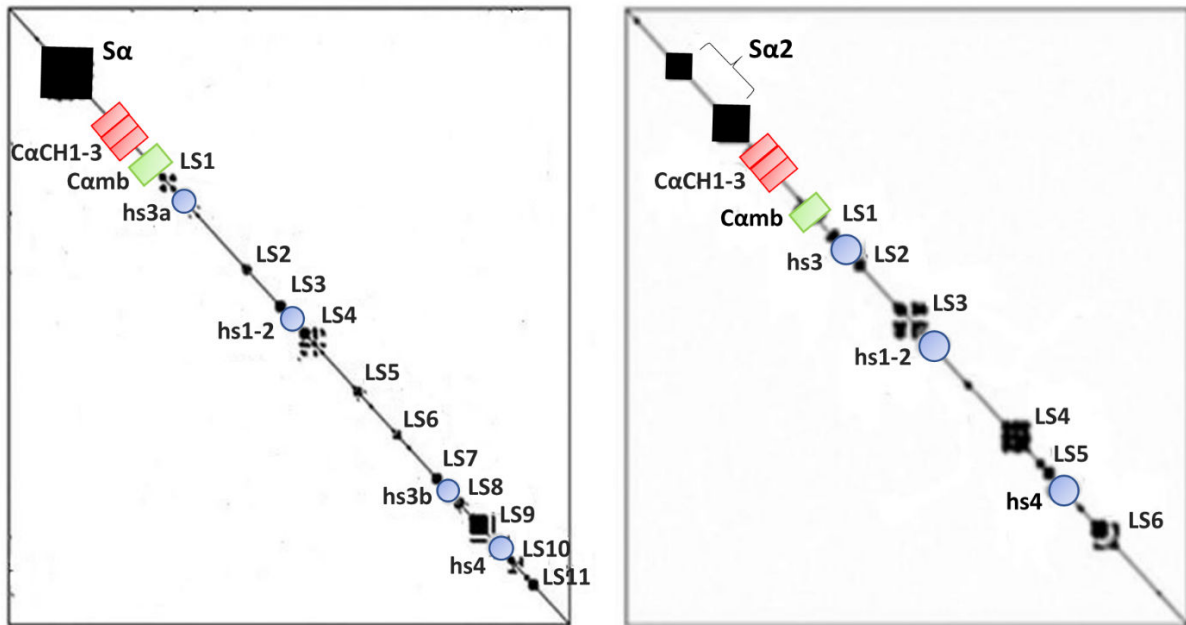


Figure 17: Localisations des LS dans la région 3'RR murin et humain

Dot plot montrant les positions des régions LS au niveau de la 3'RR chez la souris (à gauche) et chez l'Homme (3'RR2) (à droite) (Adaptée de Péron et al., 2012a).

Outre leur structure, une caractéristique majeure des régions S est la transcription germinale avant la CSR qui permet le recrutement d'AID sur l'ADN. Péron et ses collègues ont donc évalué la transcription inattendue de la 3'RR chez la souris en étudiant la transcription d'hs1-2 et d'hs4. Les transcrits sont à peine détectables dans les cellules B au repos mais apparaissent facilement lors de l'activation par le LPS.

Alors que l'accessibilité à la recombinaison LSR est sans doute assurée comme pour la CSR par la transcription de la 3'RR, la présence des régions contenant des séquences répétitives (LS) favoriserait l'accès d'AID à la 3'RR. Ce type de recombinaison entre $S\mu$ et LS (sans doute 10 à 100 fois plus rare que la CSR, mais avec une fréquence difficile à comparer puisque la CSR est associée à une survie de cellules mémoire alors que la LSR est associée à la mort cellulaire) a été détecté *in-vivo* et *in-vitro* par séquençage Sanger et par Southern Blot. De plus, une des conséquences attendues de la LSR après immunisation *in vivo* est la présence d'une petite proportion de cellules B perdant l'expression de BCR qui s'engagent dans l'apoptose (**Figure 18**) (Péron et al., 2012a, 2012b).

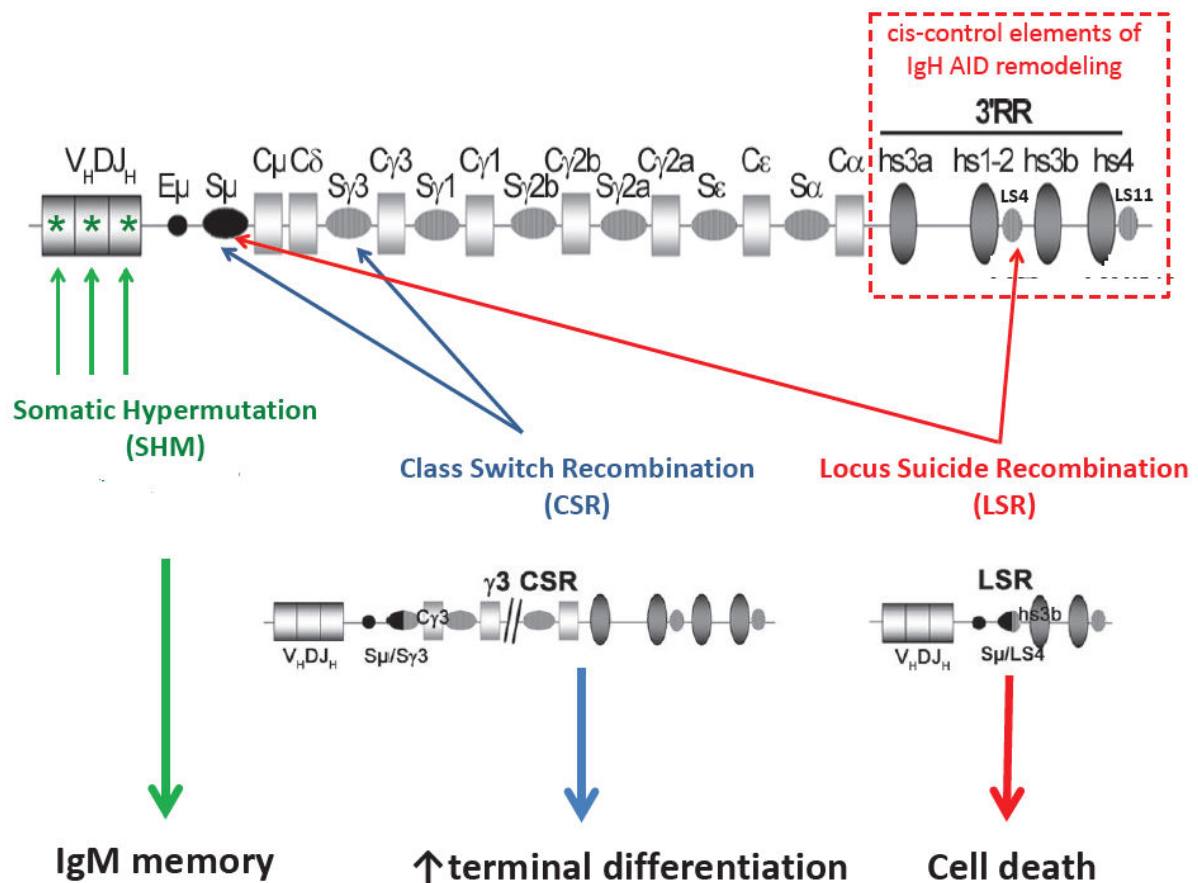


Figure 18: Mécanisme de recombinaison LSR

Lors de l'activation des lymphocytes B, l'expression de l'AID induit des mutations dans la région variable de l'Ig par SHM générant des BCR d'affinité améliorée pour l'antigène. Par la suite les cellules B changent la classe du BCR exprimé à la surface des cellules par CSR. La recombinaison suicide du locus IgH (LSR) rejoint éventuellement la S_{μ} avec l'une des régions répétitives LS situées dans la région régulatrice 3' (3'RR), puis supprime tous les gènes constants du locus IgH, conduisant ainsi à la mort des cellules B par perte du signal tonique de survie induit par le BCR (*Adaptée de la revue: Laffleur et al., 2014a*).

Au sein du centre germinatif, la LSR pourrait participer à l'homéostasie des cellules B mais elle pourrait aussi éliminer les cellules auto-réactives et les moins affines pour l'antigène. Après avoir subi la SHM, les cellules B qui ont une haute affinité pour l'antigène exprimé sur les cellules FDC bénéficierait de la CSR et d'une possible différenciation « mémoire » alors que les cellules les moins affines et qui n'ont pas reçu assez de signaux d'activation seraient éliminées par LSR ? Des études supplémentaires sont clairement nécessaires pour mieux comprendre ce phénomène, surtout chez l'Homme. J'ai ainsi analysé la survenue de la LSR chez l'Homme et la nature des jonctions observées afin de tenter d'en comprendre les modalités pendant mes quatre années de thèse.

Chapitre II : Complexe Médiateur au niveau du locus IgH

La recombinaison de classe impliquerait comme évoqué plus haut la formation d'une boucle (et/ou d'un rapprochement inter-allélique) permettant la juxtaposition des gènes entre lesquels va survenir la recombinaison de *switch*. Bien que ces contacts physiques préalables soient démontrés, les facteurs qui contrôlent ces interactions sont incomplètement connus (Mehra et al., 2016). Le complexe Médiateur est considéré comme un « multiplex » qui joue un rôle crucial dans les interactions géniques à longue distance et notamment dans celles survenant au sein du locus IgH, mais il est également impliqué dans la transcription germinale de la région *switch* acceptrice ; deux phénomènes requis pour le mécanisme de la CSR (Thomas-Claudepierre et al., 2016).

Le complexe Médiateur a été décrit en premier chez la levure *Saccharomyces cerevisiae* comme un coactivateur transcriptionnel (Kelleher et al., 1990; Kim et al., 1994). Depuis sa découverte, ce complexe s'est avéré être un composant hautement évolutif et conservé chez tous les eucaryotes (Conaway and Conaway, 2011). Sa fonction la plus connue est de réguler la transcription des gènes réalisée par l'ARN polymérase II. Bien que les mécanismes par lesquels le complexe Médiateur régule la transcription soient souvent compliqués, plusieurs études ont montré qu'il joue un rôle central dans des nombreuses voies biologiques (Soutourina et al., 2011; Youn et al., 2016). La participation du complexe médiateur dans la formation d'une boucle entre le promoteur et les *enhancers* transcriptionnels est largement documentée (Kagey et al., 2010).

Au cours de ma thèse je me suis particulièrement intéressée à étudier le rôle de ce complexe sur le locus IgH. Je détaillerai plus loin son implication dans la transcription mais aussi les recherches qui expliquent son importance sur le mécanisme de *switch*.

1. Structure du complexe Médiateur

Le complexe médiateur est composé de 25 à 30 sous-unités organisées en trois modules de base dépourvues d'une activité enzymatique : la tête, le centre et la queue ; et d'un module CDK8 (*Cyclin-Dependent Kinase 8*) supplémentaire possédant une activité kinase. Les composants de chaque module sont supposés structurellement et fonctionnellement connectés (**Figure 19**) (Tsai et al., 2014).

Les trois modules de base présentent une forte fonction activatrice de la transcription alors que le module CDK8, capable de s'associer de manière réversible aux autres modules du complexe,

a plutôt une fonction répressive en perturbant l'association du complexe Médiateur avec l'ARN polymérase II. Il est supposé que le module CDK8, formé de quatre sous-unités (CDK8, la cycline C, MED12 et MED13), fonctionne comme une unité autonome et peut exister en tant que sous-complexe. Pour réguler négativement l'expression des gènes, ce sous-complexe doit être associé au complexe Médiateur (Hengartner et al., 1998; Knuesel et al., 2009; Tsai et al., 2013).

Le module « Tête » interagit avec l'ARN Pol II, c'est le module le plus conservé pendant l'évolution et dont la perturbation libère les deux autres modules (« centre » et « queue ») du promoteur, provoquant alors un arrêt de la transcription. Le module « queue » est la structure la moins conservée et il interagit avec les régulateurs d'expression génique. Le module « centre » permet de stabiliser l'association des modules du complexe Médiateur fonctionnant comme un régulateur des interactions inter-modules (Napoli et al., 2012).

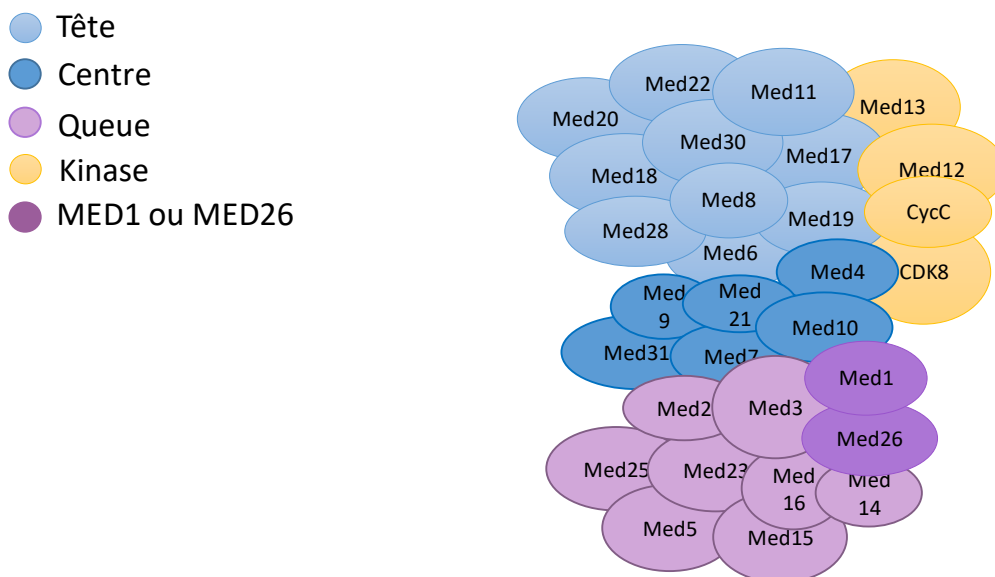


Figure 19: Représentation schématique du complexe Médiateur humain

Il est constitué de 30 sous-unités organisées en 4 modules distincts : la tête, le centre, la queue et le module CDK8. Les deux sous-unités MED1 et MED 26 ne sont pas présentes dans les complexes et sont localisés entre le module du centre et le module de la queue.

2. Fonctions du complexe Médiateur

Chez les eucaryotes, tous les gènes codant pour les protéines sont transcrits par l'ARN polymérase II (Pol II). La régulation de la transcription par Pol II nécessite l'intervention de divers activateurs et répresseurs de la transcription (collectivement connus comme facteurs de transcription spécifiques), qui se lient à leurs séquences cibles dans les régions de régulation transcriptionnelle permettant - soit par interaction directe entre les protéines ou par des actions

sur la chromatine (remodelage de la structure, modification des histones, etc.) - l'assemblage de la machinerie de transcription. Pour initier la transcription à partir d'un promoteur, la Pol II a besoin d'un ensemble de plusieurs facteurs de transcription généraux GTFs (*General Transcription Factors*) et du complexe Médiateur (Soutourina, 2018) pour former le complexe de pré-initiation de la transcription ou PIC (*Pre-Initiation Complex*) dont le complexe Médiateur a été en 1990 le dernier constituant découvert (Kelleher et al., 1990). Initialement, le complexe Médiateur était seulement considéré comme un pont protéique qui favorisait l'interaction entre les promoteurs et les *enhancers*. Des études ultérieures ont ajouté de nombreuses autres fonctions à ce complexe. En effet, des découvertes récentes montrent qu'il influence presque toutes les étapes de la transcription et coordonne ces événements en provoquant des changements de l'organisation de la chromatine (Carlsten et al., 2013).

a. Implication du complexe Médiateur dans la transcription

Pour réguler la transcription, le complexe Médiateur est recruté sur les *enhancers* des gènes. Ensuite, ce complexe interagit avec les composants du PIC positionnés sur les promoteurs qui peuvent être situés à différentes distances des régions activatrices. La régulation de son recrutement dans ces régions chromatiniennes est encore mal comprise (Soutourina, 2018).

Le premier GTF qui se fixe au promoteur est TFIID (*Transcription Factor IID*). Ensuite TFIIA et TFIIB sont recrutés suivi par le Pol II en association avec TFIIIF. Enfin, TFIIE et TFIIH complètent la formation du complexe d'initiation de la transcription PIC (Eyboulet et al., 2015). Une analyse de la localisation du PIC à travers le génome de la levure a permis l'identification des éléments de type TATA (boîte TATA) qui lient la Pol II et les GTFs (Rhee and Pugh, 2012).

Une fonction de base du Médiateur est de communiquer des signaux de régulation des facteurs de transcription liés à l'ADN directement à l'enzyme ARN Pol II. Les mécanismes précis par lesquels le Médiateur régule l'activité du Pol II restent mal compris, mais ils impliquent clairement des interactions protéine-protéine entre le Médiateur, la Pol II, et les GTFs. Deux fonctions clés supplémentaires du complexe Médiateur sont aussi décrites, à savoir qu'il stabilise le complexe d'initiation PIC en coopérant avec TFIIB, TFIID et TFIIH et il stimule l'activité kinase de TFIIH qui phosphoryle le domaine carboxy-terminal (CTD) de Pol II (**Figure 20**) (Allen and Taatjes, 2015; Soutourina, 2018).

La phosphorylation de la sérine 5 du CTD perturbe l'interaction entre le Médiateur et le Pol II ce qui permet la dissociation de Pol II du complexe PIC et la transition de l'étape de l'initiation vers l'étape de l'élongation de l'ARN (Wong et al., 2014). Le Médiateur est aussi impliqué

dans l'étape de l'élongation et de la terminaison de la transcription en recrutant des facteurs impliqués dans ces deux étapes (Mukundan and Ansari, 2011; Takahashi et al., 2011; Wang et al., 2013). Mukundan et ses collègues ont montré que le complexe Médiateur peut se positionner aux deux extrémités 5' et 3' des gènes *INO1* (*Inositol-3-phosphate synthase*) et *CHA1* (*Catabolic (L-serine) (L-threonine) deaminase*) permettant de recruter des facteurs de terminaison de la transcription lorsqu'il est positionné à l'extrémité 3' du gène (Mukundan and Ansari, 2011).

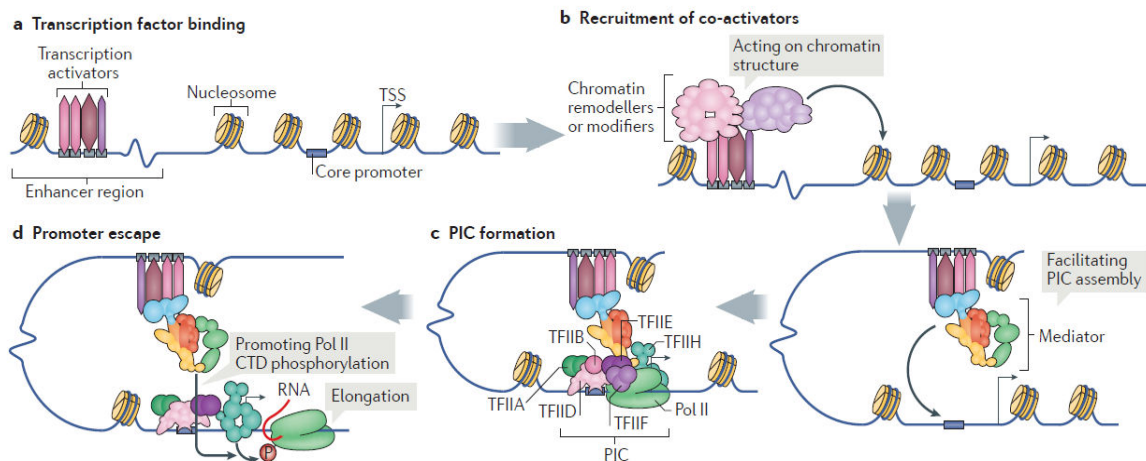


Figure 20: Activation de la transcription par le complexe Médiateur

L'activation de la transcription commence par la liaison des facteurs de transcription sur les *enhancers* (ici situées en amont du gène à transcrire). Le site de démarrage de la transcription (TSS) est indiqué par une flèche (a). Les activateurs recrutent alors des complexes de coactivateurs agissant comme modificateurs de la chromatine afin de la rendre plus accessible pour d'autres facteurs qui permettent l'assemblage de la machinerie transcriptionnelle (le complexe de pré-initiation PIC y compris le complexe Médiateur) (b). Le PIC assemblé sur le promoteur comprend le Pol II, les facteurs de transcription généraux GTFs. Le Médiateur agit pour faciliter le recrutement et / ou la stabilisation de différents composants du PIC (c). La phosphorylation de la sérine 5 du domaine carboxy-terminal (CTD) de Pol II médiée par le Médiateur est nécessaire pour que le Pol II s'échappe du promoteur permettant ainsi la transition de l'étape d'initiation à l'étape d'élongation (d) (Extrait de revue ; Soutourina, 2018).

b. Répression de la transcription par le Médiateur

Lors de l'initiation de la transcription, la Pol II qui quitte le promoteur laisse derrière elle le complexe PIC qui peut ensuite recruter une deuxième enzyme Pol II (et ainsi de suite) pour compléter plusieurs tours de transcription activée. Un tel moyen efficace de ré-initiation est probablement critique pour une expression rapide d'un gène mais il est aussi important que la transcription activée soit arrêtée pour éviter une surexpression inappropriée des gènes activés pouvant déclencher certaines maladies graves (Knuesel et al., 2009). Cette régulation est assurée par le module CDK8 qui peut avoir un impact direct et indirect sur l'activité du complexe Médiateur, Pol II et sur TFIIF (Hengartner et al., 1998).

L'arrêt de l'activité du complexe Médiateur de base est assuré principalement par la fixation du module CDK8. L'association des sous-modules Médiateur-CDK8 seule semble être suffisante pour bloquer l'interaction du Médiateur avec Pol II inhibant ainsi la transcription indépendamment de l'activité kinase (**Figure 21**) (Elmlund et al., 2006).

D'autres études ont montré que le complexe CDK8 peut phosphoryler la cycline H, ce qui inhibe TFIIH et empêche la transition de l'étape d'initiation à l'étape de l'élongation de la transcription par une inhibition de la phosphorylation de CTD du Pol II (Akoulitchev et al., 2000). Il peut aussi phosphoryler directement la CTD de l'ARN Pol II, ce qui perturbe l'association Médiateur-ARN Pol II (qui requière une CTD hypo-phosphorylée pour la fixation du Médiateur et l'assemblage du PIC) ce qui régule négativement la transcription (Hengartner et al., 1998).

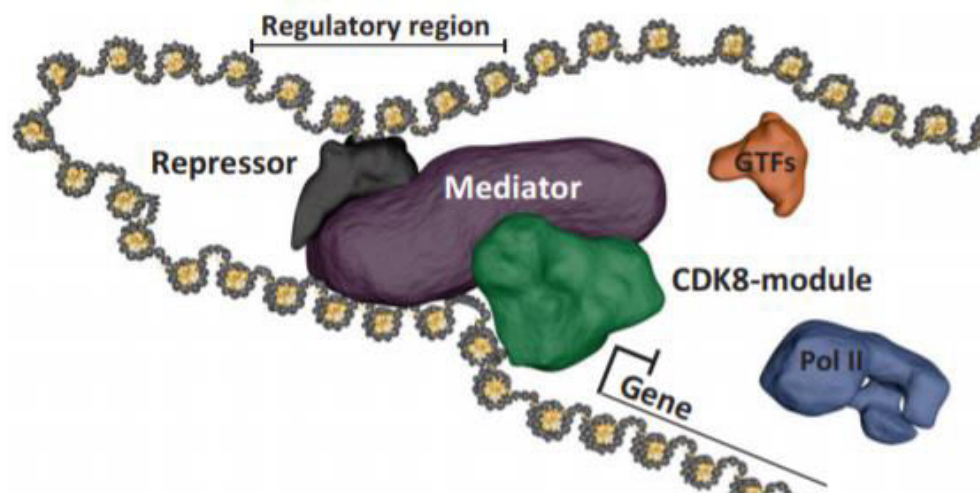


Figure 21: Inhibition de la transcription par le module CDK8

Le module CDK8 (kinase cycline-dépendant 8) (vert) peut réguler négativement les interactions Pol II-Médiateur en inhibant la transcription (*Extrait de la revue: Carlsten et al., 2013*).

c. Formation d'une boucle durant la transcription

La régulation de l'expression génique implique fréquemment la liaison des facteurs de transcription particuliers sur des *enhancers* situés à des distances éloignées du promoteur (pouvant se situer à un million de paires de base de leur promoteur chez les mammifères). Les *enhancers* sont des séquences d'ADN distales qui peuvent augmenter la transcription de leur (s) gène (s) cible (s) (éléments en *cis*). Ils peuvent être situés en amont ou en aval de leurs gènes cibles ou même placés dans les introns des gènes. Afin de stimuler la transcription, les activateurs liés à l'*enhancer* et la machinerie de transcription liée au site du promoteur doivent

être réunis, vraisemblablement par une boucle qui relie ces deux régions. Pour que ces activateurs puissent exercer leurs fonctions, la formation d'une boucle semble être logique pour rapprocher l'*enhancer* de son promoteur pouvant augmenter la concentration locale des composants de la machinerie de transcription au voisinage du gène cible où l'*enhancer* peut servir à "délivrer" le PIC à un promoteur (Cuylen and Haering, 2010 ; Pennacchio et al., 2013). La manière précise dont ces boucles sont formées reste incertaine, mais il semble que l'interaction du complexe Médiateur avec les facteurs de transcription joue un rôle important. La stabilisation de la boucle est assurée par certaines protéines dont les cohésines. Le complexe cohésine forme une structure en anneau qui engloberait l'ADN en configuration de boucle pour maintenir les interactions promoteur-*enhancer* (**Figure 22**) (Kagey et al., 2010).

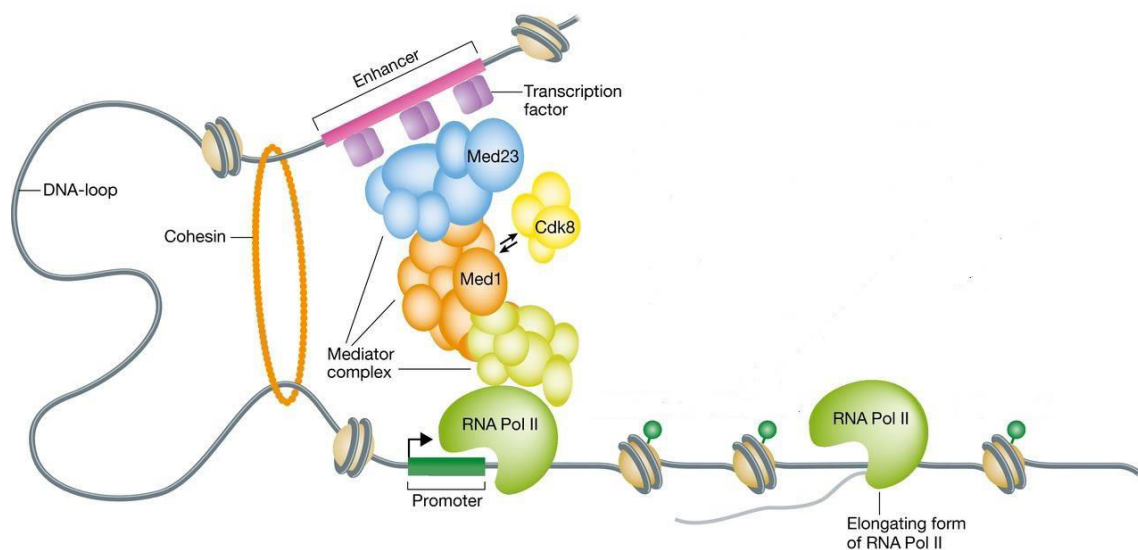


Figure 22: Formation de la boucle d'ADN par le complexe Médiateur

Le complexe Médiateur avec l'aide de la cohésine établit une boucle de chromatine, amenant la région activatrice à proximité du promoteur. Le médiateur relie ainsi les facteurs de transcription liés à l'*enhancer* aux GTF et à la Pol II (*Adaptée de la revue: Streubel and Bracken, 2015*).

d. Rôle du Médiateur dans les cellules B

Comme déjà expliqué dans le Chapitre I, au cours de la CSR le locus IgH subit des changements tridimensionnels permettant la juxtaposition des régions E μ , de la 3'RR et des régions *switch*. Les mécanismes moléculaires qui contrôlent ces changements conformationnels restent à être élucidés. La CSR implique aussi la transcription germinale des régions *switch* donneuse et acceptrice pour substituer l'isotype de l'immunoglobuline exprimé par les cellules B. Bien que les protéines qui sont impliquées dans la formation d'une boucle durant la CSR soient mal comprises, une étude récente a montré que la délétion du facteur de transcription YY1 entraîne

une baisse significative de la CSR et altère la formation de la boucle entre $E\mu$ et la 3'RR (Mehra et al., 2016). On notera à ce propos que les *enhancers* de la 3'RR portent des sites de liaison de YY1 (Pinaud et al., 2011).

L'implication du complexe médiateur dans la formation de boucles du locus IgH et dans la CSR a été étudiée par Thomas-Claudepierre *et al* qui ont montré par la technique du ChIP (*Chromatin Immunoprecipitation*) en utilisant des anticorps spécifiques pour les deux sous-unités MED1 et MED12, le recrutement du complexe Médiateur à l'enhancer $E\mu$ et aux éléments $hs1-2$ et $hs4$ de la 3'RR dans les cellules B au repos mais aussi dans les promoteurs de $\gamma1$ et de ϵ après stimulation avec du LPS+IL4 (CSR vers IgG1 et IgE) et dans les promoteurs $\gamma3$ et $\gamma2b$ après stimulation avec du LPS (CSR vers IgG3 et IgG2b). La délétion de ces deux sous-unités par des shRNA spécifiques dans la lignée CH12 (pouvant *switcher* vers IgA après stimulation *in-vitro*) entraîne une réduction significative de l'efficacité de la CSR (37%), accompagnée par une réduction de la transcription germinale de la région S acceptrice ($I\alpha-C\alpha$) alors que la transcription germinale de la région S donneuse ($I\mu-C\mu$) n'est pas affectée.

L'inactivation de MED1 *in-vivo* est aussi associée à une CSR défectueuse avec une réduction entre 30 à 60% vers tous les isotypes alors que l'expression d'AID dans ces souris n'est pas affectée, montrant un effet direct de MED1 sur la CSR. Cette délétion est accompagnée par une diminution de la transcription germinale des régions switch acceptrices et une augmentation des transcrits germinaux $I\mu-C\mu$ qui s'accumulent lorsque la CSR ne s'accomplit pas. Cette même étude a montré que la délétion de MED1 affecte également la structure tridimensionnelle du locus IgH en perturbant l'interaction entre $E\mu$ et $S\gamma1$ après stimulation *in-vitro* avec du LPS+IL4. Le gène $C\gamma1$ n'est pas efficacement mis en proximité de $S\mu$ dans les cellules B déficientes en MED1. De manière similaire, les interactions $E\mu-S\gamma2b$ induites par le LPS sont aussi moins fréquentes dans ces cellules. L'ensemble de ces données montrent que MED1 pourrait renforcer les interactions à longue distance entre les *enhancers* du locus IgH et les régions S nécessaires pour activer la transcription et l'induction de la CSR (Thomas-Claudepierre et al., 2016). La boucle formée durant la CSR est maintenue grâce au complexe cohésine (Thomas-Claudepierre et al., 2013).

La délétion de MED1 chez ces mêmes souris ne montre pas une différence significative dans la fréquence de mutation au niveau des gènes de JH4 ni dans le recrutement d'AID suggérant que le mécanisme de SHM est indépendant de la présence de la sous-unité MED1 au plan

quantitatif, même si le pattern des bases ciblées par les mutations se trouve modifié en condition de déficit Med1 (moins de mutation des paires A/T). (Gaudot et al., 2018).

3. Implication du complexe Médiateur en pathologie

Après la découverte de ce complexe, beaucoup de recherches ont montré son importance dans la régulation de la transcription à plusieurs niveaux permettant de réguler des voies de signalisation impliquées dans la croissance cellulaire, la différenciation et le développement tissulaire. Des mutations affectant des sous-unités de ce complexe sont aussi observées lors de certaines maladies et anomalies (Yin and Wang, 2014). Des altérations de l'expression de la sous-unité MED1 chez l'Homme sont associées à des dysfonctionnements cardiaques majeurs. La délétion génétique de la sous-unité MED1 chez la souris entraîne la mort embryonnaire, due en grande partie à un développement cardiaque altéré mais elle entraîne aussi des changements importants dans la fonction cardiaque comprenant la dilatation ventriculaire gauche et la diminution de la fraction d'éjection due à un changement du profil transcriptionnel du cœur (Spitler et al., 2017). Des mutations de la sous-unité MED17 du complexe Médiateur sont associées à une fréquence élevée de microcéphalie primaire caractérisée par des nombreux troubles neurologiques (Kaufmann et al., 2010).

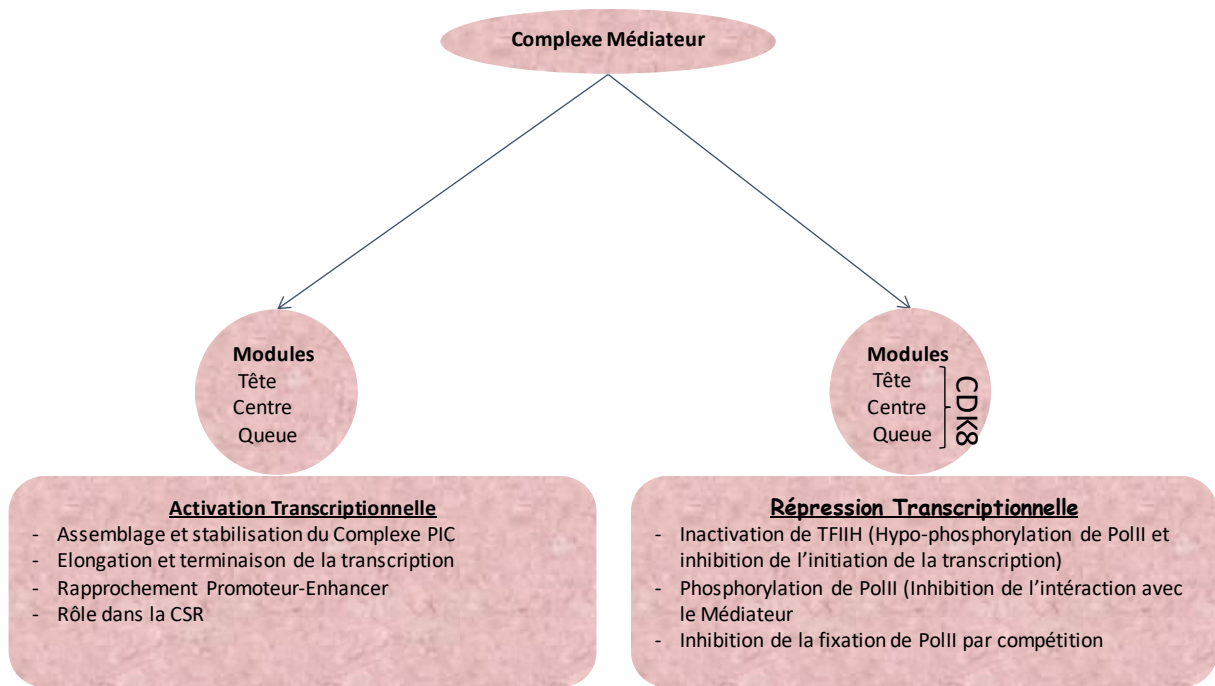


Figure 23: Récapitulatif des fonctions du complexe Médiateur

Objectifs et problématiques

Objectifs et problématiques

Au cours de leur développement, les cellules B subissent des remaniements consécutifs et majeurs du locus IgH. Lors de la phase précoce, les recombinaisons V(D)J permettent de créer la diversité du répertoire antigénique indépendamment de la rencontre avec un antigène. Durant la phase tardive de la lymphogenèse B et sous l'action d'AID, la SHM permettent la création des immunoglobulines (Ig) plus affines contre un antigène alors que la CSR permet à la cellule B et aux Ig d'acquérir de nouvelles fonctions effectrices. Enfin les cellules engagées à devenir des plasmocytes subissent une hyper-transcription des gènes d'immunoglobuline pour assurer une production élevée des anticorps (Ac).

Ces événements de la phase tardive du développement B nécessitent des interactions à longue distance au sein du locus IgH et ils sont étroitement contrôlés par des éléments *cis*-régulateurs qui contrôlent la transcription, le remodelage de la chromatine et l'accessibilité des recombinaisons, tout particulièrement le *superenhancer* 3'RR situé à l'extrémité 3' du locus. En 2012, un nouveau type de recombinaison génétique (entre S μ et la 3'RR) est décrit chez la souris au niveau de la région 3'RR. Ces recombinaisons dénommées « recombinaisons suicides du locus » (« LSR ») emportent avec elles une partie majeure du locus de chaînes lourdes d'Ig et interdisent donc aux cellules B l'expression du BCR nécessaire pour leur survie (Péron et al., 2012a).

L'objectif principal de ma thèse a été de mettre en évidence ce même type d'interactions à longue distance et de recombinaison au niveau des deux 3'RR du locus IgH humain (**Article 1**).

Afin d'aborder la mécanistique de ces recombinaisons et des interactions à longue distance qui les précèdent, j'ai également étudié le rôle du complexe médiateur sur le locus IgH et sur les remaniements qu'il subit sous le contrôle du *super-enhancer* 3'RR (**Article 2**) mais aussi sur les recombinaisons VDJ durant la phase précoce du développement B (**Article 3**) en travaillant sur un modèle murin déficient en la sous-unité MED1 dont sa délétion a un impact clair sur la physiologie du locus IgH.

Parallèlement à ces travaux, j'ai contribué à différents projets du labo qui consistaient à développer un outil bio-informatique qui permet d'analyser les jonctions CSR chez les souris mais que nous avons aussi utilisé pour analyser les jonctions CSR et LSR chez l'Homme (**Annexe 1**). J'ai aussi contribué à d'autres projet qui étudiaient l'effet de certains agents

pharmacologiques sur la CSR comme le RHPS4, un agent liant l'ADN G4, qui s'est avéré diminuer le switch *in-vitro* (**Annexe 2**).

Résultats

Article 1 : Article soumis dans Plos Research et accepté dans BioRxiv

Locus Suicide Recombination of the functionally rearranged IgH allele actively occurs in B-cells from inflamed human lymphoid tissues

Iman Dalloul, François Boyer, Zeinab Dalloul, Amandine Pignarre, Gersende Lacombe, Thierry Fest, Fabrice Chatonnet, Céline Delaloy, Anne Durandy, Robin Jeannet, Emilie Lereclus, Hend Boutouil, Jean-Claude Aldigier, Sophie Peron, Jeanne Cook-Moreau, Michel Cogné.

L'activation des cellules B entraîne une mort cellulaire abondante parallèlement à l'amplification clonale et au remodelage des gènes d'immunoglobuline (Ig) par la désamination d'une cytosine induite par AID « Activation-Induced cytidine Deaminase ». AID permet la maturation d'affinité des régions variables d'Ig et la recombinaison de classe (CSR) dans les lymphocytes B matures. Dans le locus IgH, ces processus sont contrôlés par la région régulatrice 3' (3'RR), une région qui chez la souris est à la fois transcrite et elle-même ciblée par recombinaison médiée par AID. Alternativement à la CSR, les délétions joignant $\Sigma\mu$ à la région « Like-Switch » flanquant les activateurs de la 3'RR (selon un mécanisme décrit premièrement en 2012 chez la souris) peuvent accomplir ce qu'on appelle la recombinaison suicide du locus IgH ou LSR « Locus Suicide Recombination ».

Nous nous sommes intéressés à chercher si ce même mécanisme de recombinaison suicide se produit aussi dans les cellules B humaines. Afin d'avoir accès à différents échantillons nous avons collaboré avec plusieurs laboratoires en France s'intéressant à l'activation B ou aux déficits immunitaires.

Nous montrons maintenant que la LSR se produit aussi activement chez l'Homme, est médiée par AID et induit la mort cellulaire dans plusieurs conditions d'activation des cellules B. Les délétions de la LSR se produisent soit sur l'allèle IgH fonctionnel, soit sont bi-alléliques, puisqu'elles ne peuvent être détectées que lorsqu'elles sont en cours et que leur signature disparaît au stade des plasmocytes complètement différenciés mais aussi dans les cellules B mémoires du sang « au repos » (pour lesquelles une signature LSR ne réapparaît que si les cellules B mémoires sont restimulées *in vitro*). Des points de rupture hautement diversifiés sont répartis dans les deux copies du super-*enhancer* (3'RR1 et 3'RR2) ; et il semble que toutes les conditions activant la CSR *in vitro* peuvent déclencher la LSR.

Notre travail reste descriptif, tant du phénomène que de sa mécanistique moléculaire, mais sans encore donner de clé à propos de sa régulation. Une analyse *in vivo* plus détaillée, pouvant tenir compte de l'affinité pour l'antigène, et capable de préciser dans quelle zone du centre germinatif se produit préférentiellement la LSR sera clairement utile dans l'avenir. Elle impliquera cependant la mise en place de nouvelles méthodes (qui idéalement seraient adaptées au suivi de cellules uniques).

Locus Suicide Recombination actively occurs on the functionally rearranged IgH allele in B-cells from inflamed human lymphoid tissues
 --Manuscript Draft--

Manuscript Number:	
Full Title:	Locus Suicide Recombination actively occurs on the functionally rearranged IgH allele in B-cells from inflamed human lymphoid tissues
Short Title:	LSR actively targets the human IgH locus
Article Type:	Research Article
Section/Category:	General
Keywords:	
Corresponding Author:	Michel Cogné Universite de Limoges Limoges Cedex, FRANCE
Corresponding Author's Institution:	Universite de Limoges
First Author:	Michel Cogné
Order of Authors:	Michel Cogné Iman Dalloul François Boyer Zeinab Dalloul Amandine Pignarre gersende Iacombe THierry Fest fabrice chatonnet Céline Delaloy Anne Durandy robin jeannet emilie Iereclus hend boutouil jean-claude Aldigier sophie péron jeanne cook-moreau
Abstract:	B-cell activation yields abundant cell death in parallel to clonal amplification and remodeling of immunoglobulin (Ig) genes by activation-induced deaminase (AID). AID promotes affinity maturation of Ig variable regions and class switch recombination (CSR) in mature B lymphocytes. In the IgH locus, these processes are under control by the 3' regulatory region (3'RR) super-enhancer, a region demonstrated in the mouse to be both transcribed and itself targeted by AID-mediated recombination. Alternatively to CSR, IgH deletions joining S _μ to "like-switch" DNA repeats that flank the 3' super-enhancer can thus accomplish so-called "locus suicide recombination" (LSR) in mouse B-cells. We now show that AID-mediated LSR also actively occurs in humans, and provides an activation-induced cell death pathway in multiple conditions of B-cell activation. LSR deletions either focus on the functional IgH allele or are bi-allelic, since they can only be detected when they are ongoing and their signature vanishes from fully differentiated plasma cells or from "resting" blood memory B-cells, but readily reappears when such memory B-cells are re-stimulated in vitro. Highly diversified

	breakpoints are distributed either within the upstream (3'RR1) or downstream (3'RR2) copies of the IgH 3' super-enhancer and all conditions activating CSR in vitro also seem to trigger LSR.
Suggested Reviewers:	Ali Zarrin, PhD investigator, Genentech Research zarrin.ali@gene.com Expertise about AID enzyme activity in gene recombination
	Jayanta CHAUDHURI, PhD Memorial Sloan Kettering Cancer Center chaudhuj@mskcc.org Expertise about gene recombination IN B-cells
	Javier Di Noia, PhD Universite de Montreal javier.di.noia@ircm.qc.ca Expertise about AID-mediated recombination
	Frederick Alt, PhD Professor, Harvard Medical School alt@enders.tch.harvard.edu
Opposed Reviewers:	
Additional Information:	
Question	Response
<p>Financial Disclosure</p> <p>Enter a financial disclosure statement that describes the sources of funding for the work included in this submission. Review the submission guidelines for detailed requirements. View published research articles from PLOS Genetics for specific examples.</p> <p>This statement is required for submission and will appear in the published article if the submission is accepted. Please make sure it is accurate.</p>	<p>This work was supported by grants to MC from Association pour la Recherche sur le Cancer (PGA120150202338), Agence Nationale de la Recherche (ANR grant 16-CE15-0019-01), Institut National du Cancer (INCa grant #9363), Ligue Nationale contre le Cancer and Région Aquitaine-Limousin-Poitou-Charente.</p> <p>The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.</p>

Locus Suicide Recombination actively occurs on the functionally rearranged IgH allele in B-cells from inflamed human lymphoid tissues

Iman Dalloul¹, François Boyer¹, Zeinab Dalloul¹, Amandine Pignarre², Gersende Lacombe², Thierry Fest², Fabrice Chatonnet², Céline Delaloy², Anne Durandy³, Robin Jeannet¹, Emilie Lereclus¹, HEND Boutouil¹, Jean-Claude Aldigier¹, Sophie Peron¹, Jeanne Cook-Moreau^{1,4}, Michel Cogné^{1,2,4}

¹ CNRS, UMR 7276 / INSERM, U1262, Univ Limoges, 87000 Limoges, France

² INSERM U1236, Univ Rennes ; EFS Bretagne, 35000 Rennes, France

³ Institut Imagine INSERM U1163, Paris, France

⁴ MC and JCM co-directed this work

*Corresponding authors: cogne@unilim.fr, jeanne.cook-moreau@unilim.fr.

Abstract: B-cell activation yields abundant cell death in parallel to clonal amplification and remodeling of immunoglobulin (Ig) genes by activation-induced deaminase (AID). AID promotes affinity maturation of Ig variable regions and class switch recombination (CSR) in mature B lymphocytes. In the IgH locus, these processes are under control by the 3' regulatory region (3'RR) super-enhancer, a region demonstrated in the mouse to be both transcribed and itself targeted by AID-mediated recombination. Alternatively to CSR, IgH deletions joining S μ to "like-switch" DNA repeats that flank the 3' super-enhancer can thus accomplish so-called "locus suicide recombination" (LSR) in mouse B-cells. We now show that AID-mediated LSR also actively occurs in humans, and provides an activation-induced cell death pathway in multiple conditions of B-cell activation. LSR deletions either focus on the functional IgH allele or are bi-allelic, since they can only be detected when they are ongoing and their signature vanishes from fully differentiated plasma cells or from "resting" blood memory B-cells, but readily reappears when such memory B-cells are re-stimulated *in vitro*. Highly diversified breakpoints are distributed either within the upstream (3'RR1) or downstream (3'RR2) copies of the IgH 3' super-enhancer and all conditions activating CSR *in vitro* also seem to trigger LSR.

INTRODUCTION

Humoral immune responses and immunoglobulin (Ig) production rely on the selection of B-cells harboring antigen (Ag)-specific B-cell receptors (BCRs). This selection implies not only proliferation and differentiation of those cells optimally binding Ag but also elimination of the less efficient or inappropriately activated cells. The latter can be accomplished through various pathways leading to anergy, death-by-neglect or activation-induced cell death (AICD). While AICD pathways have been characterized in detail for T-cells, and notably involve FAS-induced apoptosis, they are less documented in B-cells. A major and unique feature of mature B-cells during Ag-driven responses, is their ability to reshape their genome, and more specifically Ig genes, after activation-induced deaminase (AID)-dependent modifications. Somatic hypermutation (SHM) within germinal centers (GC) can yield cells with higher affinity V domains, which preferentially capture Ag from follicular dendritic cells, undergo stimulatory cognate interactions with T follicular helper cells and are further selected for survival. In parallel, AID-dependent class switch recombination (CSR) diversifies IgH classes by joining repetitive switch (S) regions that precede the various constant (C) genes. Besides the selected winners of AID-mediated reshaping, many cells are losers or undesired responders deserving elimination, in agreement with the considerable amount of GC B-cells which have been demonstrated as actively undergoing apoptosis [1,2]. Out-of-frame or other unfavorable V region mutations might result in BCR loss and promote apoptosis while more subtle cell fate decisions will arbitrate between death, short-term or long-term survival as memory lymphocytes or plasma cells [1,3]. Such intra-GC cell fate choices are crucial since inappropriate survival or terminal differentiation of bystander cells producing useless Ig or Ig with increased affinity for self or environmental Ags might trigger auto-immunity, inflammation and disease. Since class-switched antibodies are potent actors of auto-immunity and/or hypersensitivity, means for restricting CSR and reentry of class-switched cells into SHM thus appear as necessary safeguards to keep humoral immune responses both specific and innocuous. The AID-dependent process of locus suicide recombination (LSR) reported in mouse B-cells ideally fits this necessity [4]. We now show that the human IgH 3'RR super-enhancers include sequences ideally suited as recombination targets, and that LSR actively occurs in human lymphoid B-cells.

LSR features recombination between S μ and the 3' regulatory region (3'RR) located downstream of the IgH locus, thus deleting the whole IgH constant gene cluster. The 3'RR includes several B-lineage specific enhancers (hs3, hs1-2 and hs4) which are mostly active after B-cell activation and then promote transcription and recombination [5]. Targeted mutations of the 3'RR in the mouse demonstrated its major role in SHM, its control of germline transcription

and CSR to most C genes, its booster effect on IgH gene expression in plasma cells and even its ability to promote inter-allelic transvection between IgH alleles [5–11]. The mouse 3'RR has a unique palindromic architecture with functional implications [12–15], and this structure is shared by all mammalian species studied to date, including humans [16]. In addition, the mouse 3' RR includes multiple stretches of repetitive DNA resembling S-regions albeit shorter [4,12], potentially promoting LSR through a mechanistic process similar to CSR. In B-cells, Ag stimulation with appropriate co-stimuli induces AID expression and then initiates staggered DNA nicking in repetitive S-regions, double strand breaks (DSBs) and CSR. Highly repetitive “Like switch” (LS) regions from the 3'RR might serve a similar role, ending with S μ -3'RR junctions and deletion of the entire IgH constant region gene cluster. By abrogating BCR expression, which is critical for B-cell survival, this eliminates the corresponding cells.

In humans and old-world apes, an internal duplication of the IgH locus constant gene cluster has additionally duplicated 3' regulatory regions downstream of each C α gene (3'RR1 and 3'RR2 respectively downstream of C α 1 and C α 2)[16,17]. We show that both human 3'RRs include LS-regions which are highly structured in terms of DNA repeats and which can locally be G-rich, with predicted ability to form the G-quadruplex (G4) DNA now known as the ideal substrate for AID [18,19]. Importantly, these regions also actively undergo AID-dependent recombination in activated B-cells, whatever the activation conditions assayed, both *in vivo* and *in vitro*.

MATERIALS AND METHODS

Human samples

Blood from healthy volunteers was collected, after obtaining written informed consent, from the CHU Dupuytren Hospital, Limoges, France (ANSM: 2012-A00630-430) or from the Etablissement Français du Sang (Rennes, France; French Ministry of Higher Education and Research approval: AC-2014-2315) according to the Declaration of Helsinki. Blood from septic patients was collected from the Hematology department, CHU Dupuytren. The study was accepted by the Ethics Committee of CHU Dupuytren (n° 119-2013-19). The trial was registered under ClinicalTrials.gov: NCT01995448. Tonsils were obtained from patients undergoing tonsillectomy performed in CHU de Rennes, while adenoids were obtained from CHU Dupuytren. All were operative pieces. Bone marrow aspirates obtained from patients undergoing cardiac surgery. Subjects were recruited under institutional review board approval and informed consent according to the Declaration of Helsinki. Samples from AID-/- patients (3 PBMC samples and 1 sample of sorted tonsil B-cells) were provided by Pr Anne Durandy, Hôpital Necker, Paris, France.

Activation of human B-lymphocytes *in vitro*

Human PBMC were obtained by density gradient centrifugation and human B-cells were negatively sorted from PBMC using EasySep Human B-cell Isolation Kit (StemCell) according to the manufacturer's instructions. Purified B-lymphocytes were seeded at 1×10^6 cells/ml in IMDM medium (Lonza) supplemented with 20% fetal bovine serum (Deutscher), 1% penicillin/streptomycin (Gibco) and stimulated for 4 days with 100 ng/ml human recombinant CD40L (Enzo Life Sciences) alone or with 50 ng/ml recombinant human IL-4 (Peprotech) or 2 µg/ml Gardiquimod (InvivoGen) or 2.5 µg/ml CpG oligodeoxynucleotide 2006 (InvivoGen) or 100 ng/ml IL-21 (R&D Systems) or 100ng/ml INF γ (R&D Systems) or 50 ng/ml Pam3CSK4 (InvivoGen) or 0.5 µg/ml goat anti-human kappa (Southern Biotech).

Sorting of memory, transitional, and naive B-cells, and generation of *in vitro* activated plasmablasts from cultured naive B-cells

PBMCs from healthy volunteers were obtained after density centrifugation. Naive B-cells (NBCs) were purified by negative selection using magnetic cell separation (Naive B-cell Isolation Kit II; Miltenyi Biotec), following the manufacturer's instructions, with the AutoMACS deplete-sensitive program. Purity of isolated CD19+CD27- NBCs was routinely >99%. Purified NBCs were labeled with 1 µM CFSE (Thermofisher) at 37°C for 10 min and washed in complete medium consisting of RPMI 1640 supplemented with 10% FCS and antibiotics (all from Gibco, Thermofisher). Purified human NBCs were cultured at 7.5×10^5 cells/ml in 24-well plates and stimulated during 4 days with 2.6 µg/ml F(ab')₂ goat anti-human IgA+IgG+IgM (H+L) (Jackson ImmunoResearch Laboratories, West Grove, PA), 100 ng/ml recombinant human soluble CD40L (NCI), 2 µg/ml CpG oligodeoxynucleotide 2006 (Miltenyi Biotec), and 50 U/ml recombinant IL-2 (SARL Pharmaxie). Day-4 activated B-cells were washed and cultured at 4×10^5 cells/ml with 50 U/ml IL2, 12.5 ng/ml IL-10 and 5 ng/ml IL-4 (R&D Systems). At day-6, cells were separated by

cell sorting (FACSAria cell sorter, BD Biosciences) into undifferentiated CFSE^{hi} CD38^{lo} (bystander lymphocytes, without proliferation), CFSE^{lo} CD38^{lo} (proliferating activated B-cells), and differentiated CFSE^{lo} CD38^{hi} (plasmablasts) subsets[20].

After B-cell isolation (B-cell isolation kit II, Miltenyi Biotec), we also purified transitional, naive and memory B-cells which were sorted as CD27+CD38^{lo}CD24^{lo}, CD27-CD38^{lo}CD24^{lo} and CD27-CD38^{hi}CD24^{hi} subsets, respectively. Isolated cells were stimulated 4 days with different combinations of 100 ng/ml CD40L, 50 ng/ml IL4 and 2 µg/ml Gardiquimod (InvivoGen).

Sorting of *in vivo* differentiated early plasmablasts

Actively differentiating plasmablasts from tonsils were purified as previously described [21] as CD19+IgD-CD10-CD38^{hi} cells. CD38^{hi}/CD138+ mature plasma cells were purified from bone marrow aspirates.

RNAseq on activated human B-cells

Human naive B-cells were purified and activated *in vitro* as described (Hipp et al. Nat Commun. 2017). RNA extractions were performed on day 4 IL-2 primed cells with the NucleoSpin RNA XS kit (Macherey-Nagel). Libraries were prepared with the TruSeq Stranded mRNA Library Prep Kit (Illumina) and samples were sequenced on an Illumina NextSeq 500 using 75-bp single-end reads (NextSeq 500 High Output v2, Illumina) by Helixio (Clermont Ferrand, France). Quality of sequencing data was monitored by FastQC. Residual adapters from sequencing were trimmed using Cutadapt 1.0. Potential PCR duplicates were removed using SAMtools 1.3. Reads were then aligned on the GRCh38 human genome using STAR 2.4.2a.

Amplification of S_μ/S_γ for CSR-seq experiments

DNA from B-cells was extracted using [GenElute Mammalian Genomic DNA Miniprep Kit](#) (Sigma). S_μ/S_γ junctions were amplified in triplicate by nested PCR with 200 ng DNA ([Herculase II Fusion DNA Polymerase, Agilent Genomics](#)) using the following consensus primers which amplify all four classes of human IgG junctions: PCR1: S_μ1a 5'- CCAGGTAGTGGAGGGTGGTA-3' and IgGa consensus reverse 5'- GGTCACCACGCTGCTGAG-3', PCR 2: S_μ1b forward 5'- CAGGGAAGTGGGGTATCAAG-3' and IgGb consensus reverse: 5'- CTTGACCAGGCAGCCCAG-3'. Forward primers used in both PCRs are located in S_μ while the reverse primers are located in the IgG consensus sequence of the CH1 region.

Amplification of S_μ/3'RR junctions for LSR-seq experiments

S_μ/3'RR junctions were amplified under the same conditions as for S_μ/S_γ junctions using different primers located at different positions within the 3' RR. The forward primer located in the E_μ region (E_μH1) was used with a reverse primer located downstream of Hs4 (3'FarHs4 reverse) or with a primer located in Hs1.2 (Hs1.2 reverse1) or in Hs3 (Hs3 reverse 1). PCR1 product was used to perform PCR2 using S_μ1b forward with all reverse primers located in 3' RR. The primers sequences and PCR programs are given in Table I.

Table I. Primers used in nested PCR for amplification of LSR junctions

PCR 1 Primers	Sequence
EμH1 Forward	5'- AGGCTGACCGAAACTGAAAA-3'
3'farHhs4 reverse	5'- CAAGCGTCAAGGTGTGGAC-3'
Hs1,2 reverse 1	5'-TTCCCAGGGGTCCTGTGGGTC-3'
Hs3 reverse 2	5'- CTGGCCAGGTCTCGGTTTT-3'
PCR2 primers	Sequence
Sμ 1b Forward	5'-CAGGGAAGTGGGGTATCAAG-3'
Probe 3'Hhs4 reverse	5'-GGACGCGGTTTGCTTTTAT-3'
Hs1,2 reverse 2	5'-GCTGAGTTTTTCGGCATCTCT-3'
Hs3 reverse 1	5'-GTTTTGGGGCATGTTTCTCA-3'

Ion torrent next generation sequencing NGS

Barcoded libraries with 200-pb read lengths were prepared using Ion Xpress plus Fragment Library Kit (Thermo Fisher Scientific) according to manufacturer's instructions. Each library was prepared using 200 ng PCR2 product. For LSR junction sequencing equal amounts of each PCR2 product (using hs3, hs12 and h4 primers) were mixed together and diluted to 100 pM to prepare one barcoded library. Libraries were run on an Ion PI v3 chip on the Ion Proton sequencer (Life Technologies). Data analysis was performed using CSReport [22].

Data analysis:

Data was analyzed by using GraphPad Prism software. Data are shown as means \pm SEM of the indicated number of values. The student T test was used to determine significance between stimulation conditions.

RESULTS

Human 3'RR1 and 3'RR2 share multiple common features with S regions

Architectures of human 3'RR1 and 3'RR2 elements were analyzed for the presence of direct and inverted repeats (Fig1). As previously described for the mouse 3'RR, “like-switch” (LS) sequences (*i.e.* stretches of highly repetitive DNA somehow resembling S regions) were readily evidenced using the YASS algorithm (<http://bioinfo.lifl.fr/yass/yass.php>) for a survey of direct repeats in window range set over 20bp, with a minimal identity threshold set at 90% (Fig. 1A). Five to six human LS regions stand in each of the human IgH 3'RR super-enhancers. These LS regions are each 0.5 to 2 kb long, and are interspersed with core enhancer elements, which they do not overlap.

Human 3'RR LS sequences additionally show 35-50% G-richness on one or the other DNA strand, and analysis using the G4-hunter algorithm <http://bioinformatics.cr.uk/G4Hunter/> [23], predicted the occurrence of several clusters of G4-DNA, standing in either one or the other DNA strand of LS regions (Fig 1B).

When checking RNA-seq data from activated human B-cells, we noticed that 3'RR transcription occurred in both the sense and antisense orientation in reference to IgH constant genes, and that, quantitatively, it revealed equivalent amounts of transcripts corresponding to either orientation (Fig. 1B). This is in contrast with IgH constant genes (see for example C α and the preceding S α region on Fig 1B), which are predominantly transcribed in the sense orientation.

LSR occurs at multiple sites within the whole extent of both human 3'RR

As previously described for mouse LSR, we identified human LSR junctions by nested PCR in DNA from B-cell samples [4]. The specificity of the read-out for these PCR assays was improved by replacing the final hybridization step of PCR products, initially done with specific probes[4]), with high-throughput sequencing adapted from our recently described CSR evaluation algorithm, CSReport [22]. Sequence determination thus precisely mapped the breakpoint sites. CSReport mapped LSR junctions linking S μ to either the human 3'RR1 downstream of C α 1 or 3'RR2 downstream of C α 2.

Although such an “LSR-seq” DNA amplification method (with some amplified fragments above 30kb) is necessarily biased towards amplification of the shorter fragments, the use of 3 primers located either downstream from the h3, hs1,2 and hs4 enhancers finally amplified breakpoints located throughout 3'RR1 and 3'RR2.

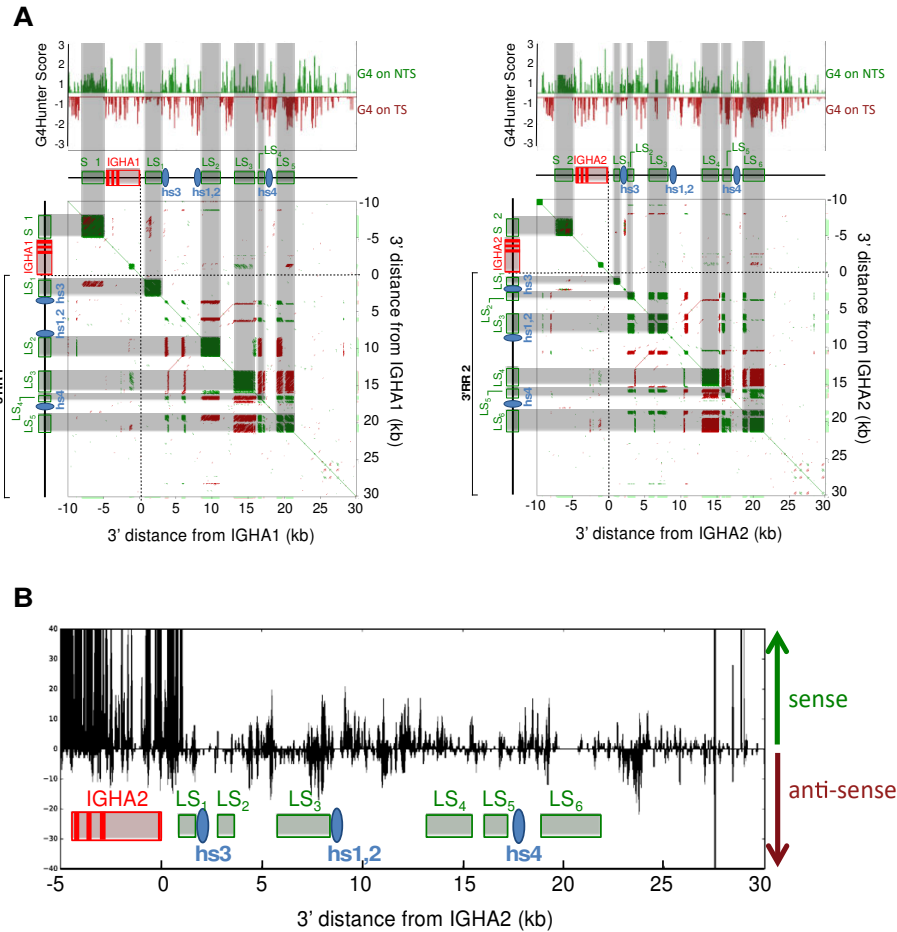


Figure 1: 3'RR structure and transcription in activated B-cells. A. Structure of genomic fragments including either human $\alpha 1$, $\text{C}\alpha 1$ gene and 3'RR1 (left), or $\alpha 2$, $\text{C}\alpha 2$ and 3'RR2 (right) was analyzed for presence of G4 DNA on the template (TS) or non-template (NTS) strand (*Top*). Dot plot analysis of internal repeats (*Bottom*) was ran with the YASS algorithm: colored dots indicate short stretches of homology in direct (green) or inverted (red) orientation (green squares thus mark positions of highly repetitive S or LS regions / red lines indicate positions of inverted repeats, notably the palindrome centered on the hs1,2 enhancer). **B.** Transcription of human 3'RR2 in sense and antisense orientation, as evaluated by RNA-seq experiments with RNA from *in vitro* activated B-cells.

These breaks were joined to various positions within $\text{S}\mu$, with preferential use of breakpoints at the 5' side of $\text{S}\mu$, especially for junctions with the downstream part of the 3'RRs. When considering the sum of all junctions mapped in this study (Fig. 2), the most striking observation was their highly diversity and scattered distribution throughout the 3'RRs (both within LS regions and in their proximity).

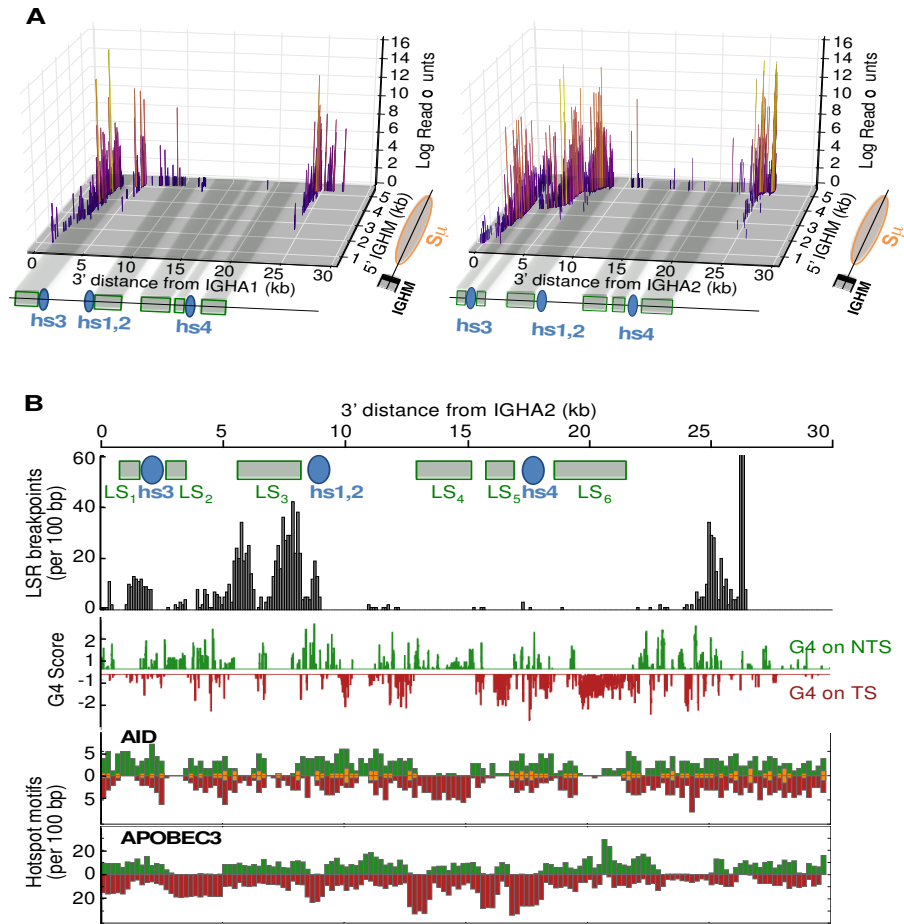


Figure 2: Positions of LSR breaks in human activated B-cells. **A.** A 3D-schematic showing the positions of junctions sequenced in this study with regards to abundance among sequence reads (in log scale) and position of DSBs in S_{μ} , and 3'RR1 (left) or 3'RR2 (right). **B.** Map of 3'RR2 reporting breaks observed in sequenced LSR junctions together with the position of G4 DNA on either the non-template (green) or template (red) strand (with reference to the transcriptional orientation of IgH constant genes), AID consensus target sites, and APOBEC3 family deaminases.

It is questionable whether LSR targets only unswitched B-cells, determining the occurrence of S_{μ} -3'RR junctions, or can also target class-switched cells and then eventually conserve a remnant of a downstream S region inserted in-between S_{μ} and a recombined 3'RR. We indeed obtained multiple evidence of such “post-switch” LSR, with sequence reads obtained from the LSR-seq protocol (*i.e.* sequencing an S_{μ} to 3'RR amplicon) but including junctions of S_{μ} with S_{γ} or S_{α} , (or S_{γ} or S_{α} joined with a 3'RR). Within a total of 11179 LSR junctions sequenced in this study, 1126, *i.e.* about 10% could be typed as complex junctions involving breaks with a downstream S (S_{γ} or S_{α}) region and certifying their origin from switched cells. Conversely, direct S_{μ} -3'RR junctions could originate from either unswitched or switched cells (upon complete deletion of the intervening S regions), so that the sequence of a given LSR junction cannot by itself certify that it occurred in an IgM+ cell.

LSR-seq can semi-quantitatively estimate *in vivo* LSR in freshly collected B-cells

CSR junctions were characterized in parallel to LSR junctions, by a similar PCR/sequencing strategy in order to provide an LSR/CSR comparison. “LSR-seq” and “CSR-seq” protocols thus provided two semi-quantitative assessments of recombination events: numbers of sequencing reads including the expected junctions, and numbers of unique junctions. Although we cannot claim such a nested PCR and sequencing assay to be fully quantitative, it appeared to be at least semi-quantitative, giving negative results in non-B-cell samples and abundant signals in B-cells from chronically inflamed lymphoid tissues. It also had the abovementioned advantage of formally identifying sequencing reads mapped as CSR or LSR junctions, and allowed us to count repeated reads with identical sequences as “unique junctions”.

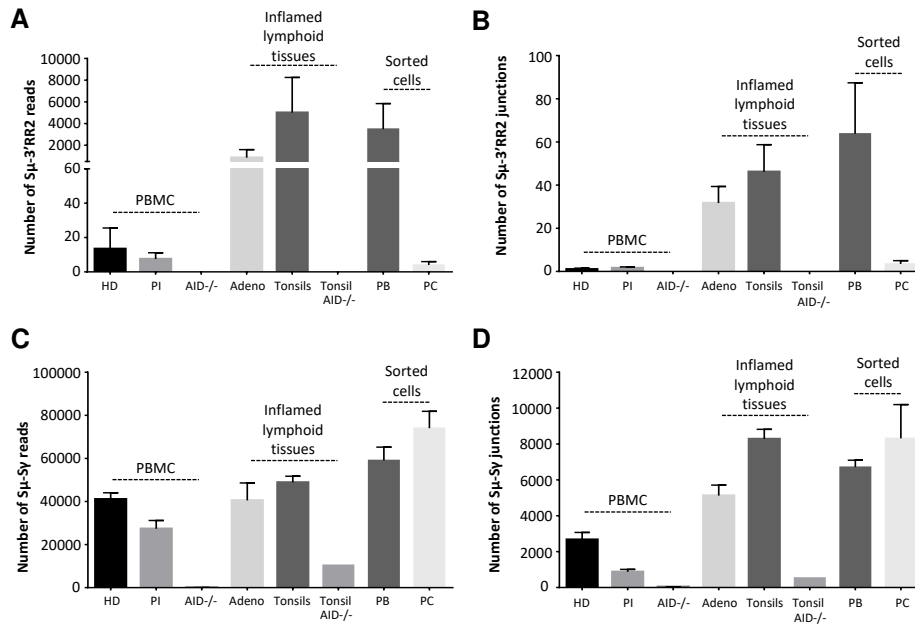


Figure 3: LSR breaks *in vivo* in humans. Sμ-3'RR2 LSR reads (A) and junctions (B) were analyzed in peripheral blood B cells isolated from healthy donors (HD) (n=6) but also in blood from post-infection (PI) patients recently hospitalized and treated for bacterial infections (n=17) and from AID-/- B-cells (n=3). LSR was also studied using DNA from inflamed lymphoid tissues (adenoids (n=3) and tonsils (n=11)), 1 AID-/- tonsil, from sorted plasmablasts (PB) isolated from tonsils (n=5) and finally in DNA from terminally differentiated plasma cells (PC) (n=2) sorted from bone marrow. Frequencies of LSR junctions amplified with an Sμ forward primer and a downstream Hs4 reverse primer were evaluated in parallel to those of classical Sμ-Sγ CSR reads (C) and junctions (D) amplified from the same DNA samples, in order to provide an LSR/CSR comparison, data represent mean numbers (reads or junctions) ± SEM.

LSR junctions from S μ to either 3'RR were accordingly "semi-quantified" in DNA from fresh primary B-cells from various origins. Peripheral blood mononuclear cells (PBMC) from healthy donors regularly yielded low, close to negative, LSR signals, and the same was observed in a series of patients recently treated for bacterial infections. By contrast, LSR was abundant in DNA from tonsils or adenoids. No LSR junctions were found in DNA from either tonsil B-cells nor PBMC obtained from AID-deficient patients. In healthy donors, as expected from the weak signals from PBMC, LSR was also absent or very low in sorted [CD19+, CD27+] peripheral blood memory B-cells and in sorted bone marrow CD138+ plasma cells (Fig. 3). LSR junctions were thus rare in DNA from circulating B-cells of healthy subjects and in differentiated plasma cells or memory cells. Further splitting and sorting of tonsil cells showed that LSR junctions occurred at a high level in cells with ongoing activation (sorted as blasting CD19+IgD-CD10-CD38^{hi} B-cells) (Fig. 3).

LSR junctions actively occur *in vitro* in activated B-cells

Class-switching and induction of AID in human B-cells is mostly a T-dependent process and is most often modelled *in vitro* by activating B-cells with CD40 ligands. When total blood B-cells from healthy patients were stimulated in such conditions, this resulted in the induction of LSR simultaneously with CSR. Associating anti-CD40 stimulation with various co-stimulatory factors, increased LSR and the LSR/CSR ratio in most of the conditions assayed and notably upon addition of IL4, IL21, or BCR ligation with anti- κ antibodies, or, to the highest extent, upon TLR ligation (Fig. 4).

We also sorted three different activated B-cell stages, using a 6-day long *in vitro* activation protocol with combined anti-CD40, TLR9 ligand CpG, anti-BCR antibodies and cytokines, previously described as optimal for mimicking the *in vivo* humoral immune response [21] (ref). This allowed us to generate either plasmablasts committed towards plasma cell differentiation, or activated B-cells having proliferated but still at an earlier stage (resembling proliferating centroblasts), and to compare them to surviving but non-dividing B-cells (bystanders) from the same culture. Analysis of the three types of cells showed that LSR was virtually absent in non-dividing bystander B-cells, but was, by contrast, at the highest level in plasmablasts and at an intermediate level in actively proliferating activated B-cells (Fig. 5A).

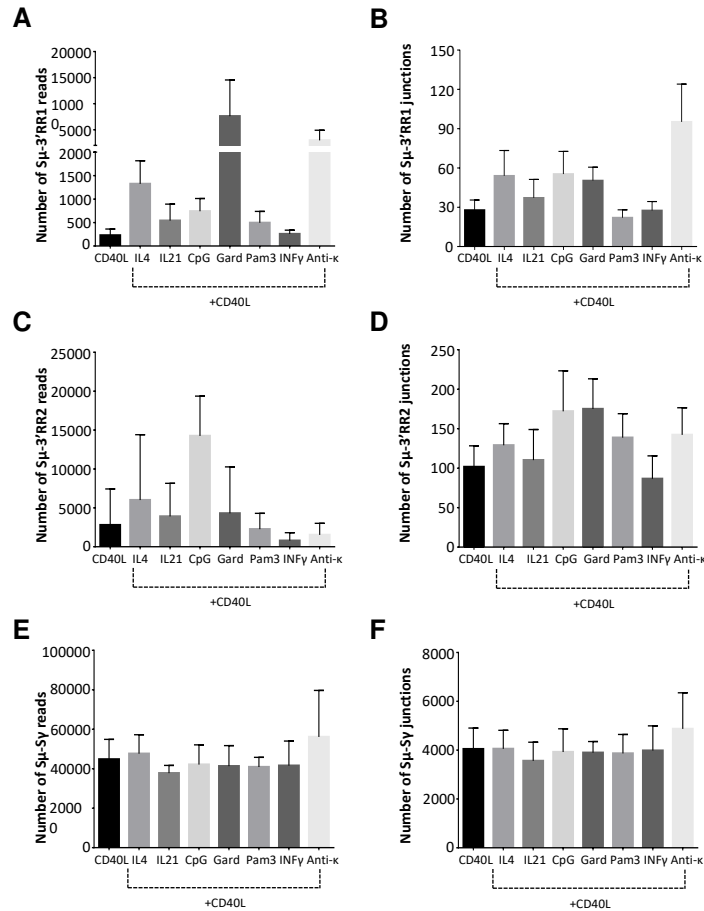


Figure 4: LSR after *in-vitro* stimulation of peripheral blood B-cells in various conditions of activation. PBMC from healthy donors activated for 4 days with human recombinant CD40L alone or with recombinant human IL4, IL21, CpG, gardiquimod (Gard), Pam3, IFN γ or with an anti-Ig κ antiserum. S μ -3RR1 (A and B) and S μ -3RR2 (C and D) LSR junctions were amplified using reverse primers located in hs3, hs1-2 and downstream hs4. Equal amounts of the three PCR products were mixed in a single library. E and F represent IgG CSR reads and junctions using the same stimulation conditions. Data represent means (for unique reads and junctions) \pm SEM for 6 people.

We also studied LSR in cells first sorted and then stimulated *in vitro*, comparing memory, naive and transitional B-cells (sorted as CD27+CD38^{lo}CD24^{lo}, CD27-CD38^{lo}CD24^{lo} and CD27-CD38^{hi}CD24^{hi} subsets, respectively). We found that upon *in vitro* activation, memory B-cells yielded much higher levels of LSR than naive or transitional B-cells (LSR remaining undetectable in control Jurkat cells) (Fig. 5, E and F). As expected in sorted populations including from the beginning a large amount of switched cells, CSR junctions assayed in parallel were the most abundant in memory cells (Fig. 5, G and H).

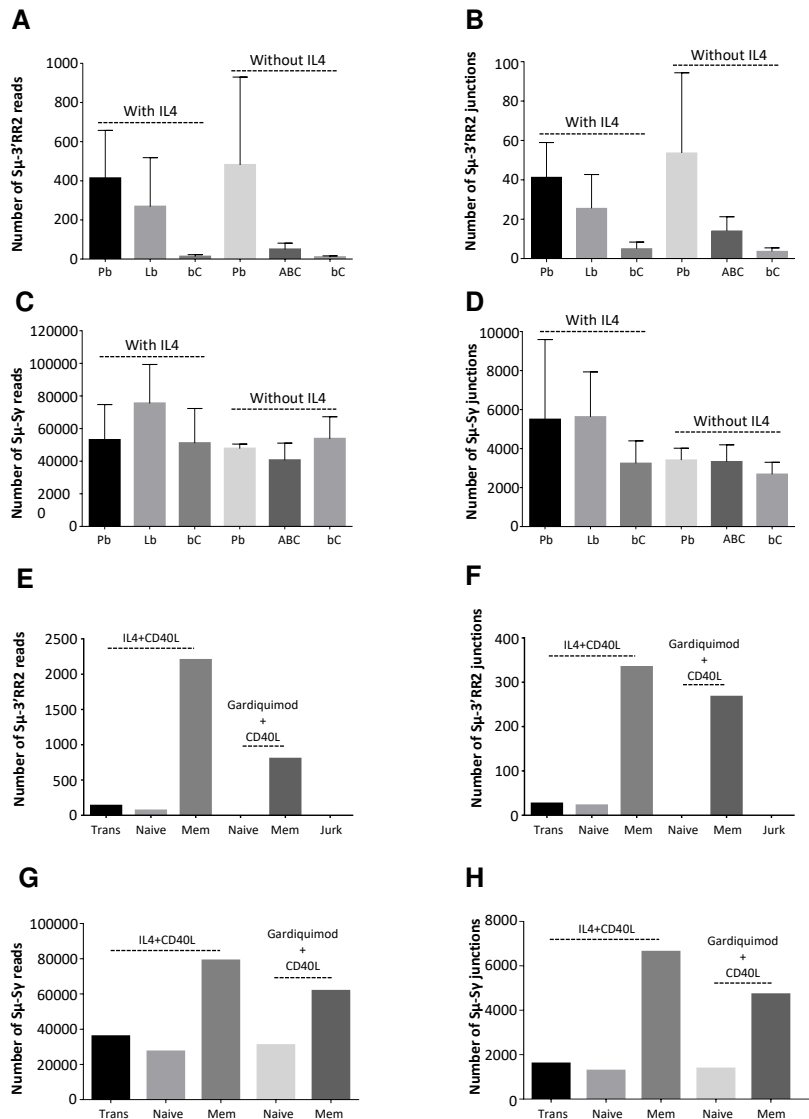


Figure 5: LSR breaks in vitro in stimulated B-cells from various origins. **A** and **B**. LSR (reads and junctions) in *in vitro* stimulated naive B-cells, sorted after stimulation. PBMC from healthy donors were activated *in-vitro* for 6 days with combined anti-CD40, TLR9 ligand CpG, anti-BCR antibodies and other cytokines generating three different stages of differentially activated- B-cells: Pb (plasmablasts), ABC (proliferating activated B-cells) and bC (bystander B-cells). Data represent mean number (reads or junctions) \pm SEM for 3 donors. **C** and **D**. CSR (reads and junctions) to Sy in same samples as A and B. CSR junctions were amplified using a downstream IgG consensus primer located in CH1. **E** and **F**. Reads and junctions from *in vitro*-induced LSR in stimulated B-cells from various origins: sorted transitional (Trans), naive and memory B-cells (Mem) (n=1) B-cells obtained from peripheral blood of healthy donors. Before DNA extraction, cells were stimulated for four days *in-vitro* with human recombinant CD40L combined with human recombinant IL4 or with gardiquimod. Jurkat (Jurk) cells stood as negative controls. **G** and **H**. CSR reads and junctions from same samples as in E and F. CSR junctions were amplified using a downstream C γ consensus primer located in CH1. Data represent means (for unique reads and junctions) \pm SEM.

Structure of LSR junctions

CSR junctions from S_{μ} to S_{γ} and LSR junctions from S_{μ} to 3'RR were amplified by PCR with specific primers, sequenced by NGS (Ion Proton) and analyzed regarding their positions in S regions and/ or 3'RR, and their nature using CSReport, a dedicated tool that performs junction assembly and analysis based on BLAST alignments (Fig. 6). A first difference between LSR and CSR concerned the global position with regard to repetitive regions. All S_{μ} breaks were located in the repetitive part of S_{μ} (in agreement with the fact that in the mouse, deletion of S_{μ} repeats virtually abrogates CSR)[24]. By contrast in the 3'RRs, breaks were found not only within but also outside of LS regions. A second difference concerned the detailed position of 3'RR breaks (as identified in sequenced unique junctions) and their increased distance relative to the various genomic features relevant to DNA cleavage (main AID target sites WRCY and RGYW, G4 sites suggested by the G4 Hunter algorithm) (Fig. 6, A and B). Interestingly in LSR junctions, the increased mean distance concerned only 3'RR breaks, while S_{μ} breaks were similar to CSR junctions. These increased mean distances between breakpoints in unique junctions and relevant genomic features (in color) were compared to the same distance calculated from simulated random break positions in the sequenced regions. Even if the mean [break to AID site] distance was longer in the 3'RR than in S_{μ} , it remained significantly shorter than that calculated under the random break hypothesis. The same evaluation was done for APOBEC3 family target sequences, which by contrast did not show any difference between the observed distance and that calculated under the hypothesis of random breaks (Fig. 6C).

Beyond DNA breaks, another difference between CSR and LSR was related to the junctional sequence itself. Our CSReport algorithm indeed classifies junctions as either blunt (broken ends directly joined without overlap), with microhomologies (stretches of a few nucleotides shared between joined segments) or with insertions (a few nucleotides inserted between joined segments). Figures 6, D-F compare more than 2000 LSR unique direct junctions characterized in this study ($\approx 20\%$ S_{μ} -3'RR1 and $\approx 80\%$ S_{μ} -3'RR2 junctions) to a thousand of independent unique S_{μ} - $S_{\gamma}1$ junctions obtained from one tonsil sample. LSR and CSR somehow differed : LSR to both S_{μ} -3'RR1 and S_{μ} -3'RR2 exhibited a lower frequency of blunt junctions (12 % vs. 19%) and longer microhomologies (mean length 3.8bp vs. 2.7bp) and insertions (mean length 4.8bp vs. 3.7bp).

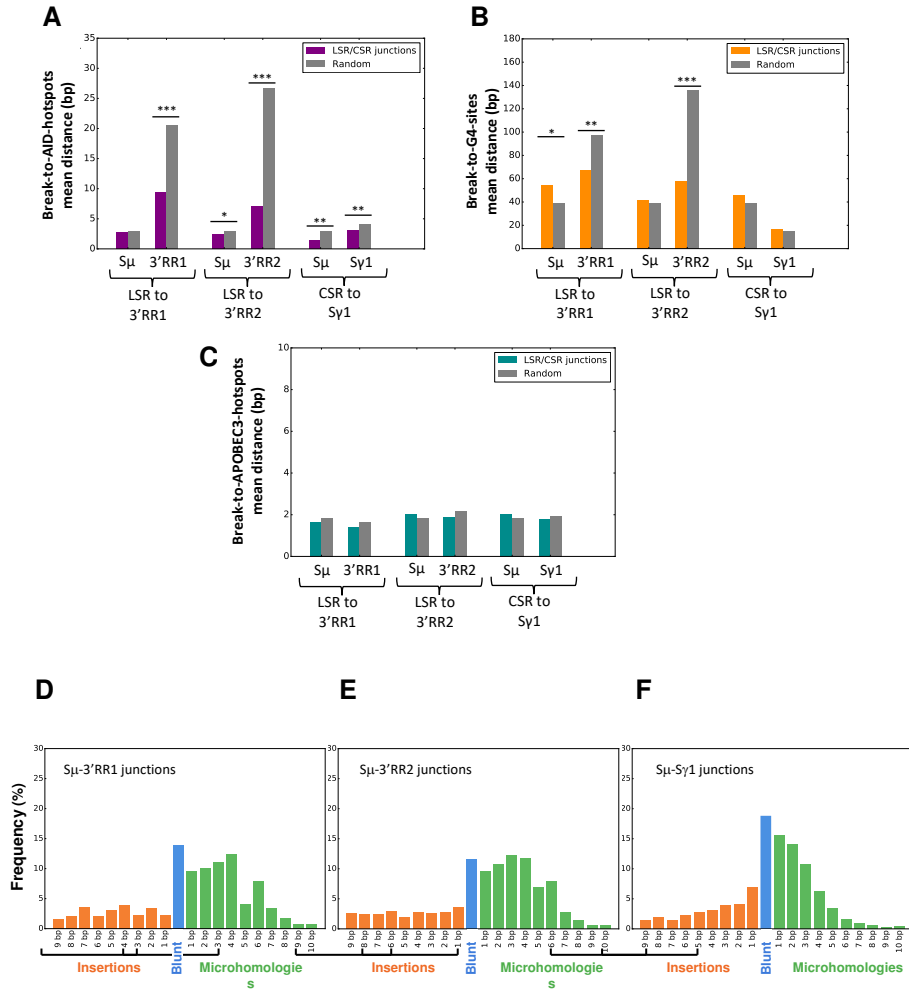


Figure 6: Detailed position of breaks and structure of human repaired LSR junctions. Mean distance to AID target sites (A), G-quadruplex (B) and APOBEC3 family (C) sites. Structure of repaired junctions with regards to mean number of inserted nucleotides and length of microhomology between broken ends for S_{μ} -3'RR1 (D), S_{μ} -3'RR2 (E) and S_{μ} -Sy1 (F).

DISCUSSION

Analyzing the architecture of human IgH 3'RR elements shows their high architectural similarity with the mouse 3'RR, notably for the presence of a central palindrome centered on the hs1,2 enhancer and for the presence of several stretches of repetitive LS regions. The latter LS regions are each 0.5 to 2 kb long, which is close to the length of *Sα2* and appears longer than LS regions in the mouse 3'RR, with the longest LS flanking the hs4 core enhancer.

Although in lower amounts than for S-regions, human 3'RR1 and 3'RR2 also include G-rich DNA, but not restricted to a single DNA strand. A specific feature of mammalian S regions is indeed G-richness of the non-template strand, where it promotes the initiation or extension of R-loops, and also often determines tandem G-repeats with potential to form G4 DNA [18,19,25]. In S regions, G4-DNA promotes the occurrence of DNA nicks and further excision by exonuclease 1, while pharmacological agents binding G4-DNA can inhibit CSR [18,26,27]. *In vitro*, G4-DNA was also shown to provide a type of structured DNA which favors the activity of Activation-Induced Deaminase (AID) [19]. Finally, G-quadruplexes forming in S region transcripts were shown to recruit the helicase Ddx1 and contribute to the formation of R-loops before CSR [28].

The G4-hunter algorithm also identified G4-DNA forming sequences within the human 3'RRs, together with long G-rich stretches standing in either one or the other DNA strand. In this regard, it should be noticed first that human 3'RR transcription occurs in both the sense and antisense orientation (with reference to *IgH* constant genes) and secondly that 3'RR eRNA are non-coding transcripts, for which definition of a "template strand" (as for *IgH* constant genes) is ambiguous. In such settings, G4 structures from any DNA strand might recruit AID activity within the 3'RR LS regions, similarly to G4-DNA from the non-template strand within S-regions.

The present study demonstrates a high diversity of IgH recombination events joining $\text{S}\mu$ or downstream S regions to the 3'RR super-enhancers in activated human B-cells from healthy individuals, showing inter-species conservation of the LSR process previously observed in the mouse. This process appears to be AID-dependent since such junctions were virtually absent in tissues from AID-deficient patients. Study of cells from various origins indicated that human LSR junctions could only be detected transiently when cells were activated but were not preserved in the long term in quiescent memory B-cells. Indeed, when DNA from peripheral blood B-cells was collected from various individuals or from bone marrow plasma cells, it gave low or negative LSR signals by LSR-seq, while CSR signals were high. Since peripheral blood is known to include one-third of memory B-cells in adults [29], this observation suggests that IgH alleles carrying LSR deletions are absent (or exquisitely rare) in circulating

memory B-cells and thus that they rarely target the non-functional allele. Rather, LSR deletion likely predominantly targets the functional IgH allele (or is biallelic), and therefore incompatible with further survival of B-cells in the absence of a functional BCR. On the contrary, a major portion of circulating memory B-cells obviously survive class-switching and accounts for persisting CSR junctions detectable in DNA from PBMC [29].

By contrast to blood, human lymphoid tissues, known as sites of chronic immune stimulation, such as tonsils and adenoids, yielded abundant LSR junctions by LSR-seq. Contrasting with spleen or blood, tonsils and adenoids are known to include abundant GC B-cells with an ongoing activation phenotype (CD20^{high}, BCR low, often showing CD38 expression) [29,30]. B-cells from such chronically activated lymphoid tissues thus include a mix of naive cells, memory cells (about 20%) and GC activated B-cells as well as plasmablasts involved in ongoing local immune reactions to mucosal antigens.

This major difference between blood and tonsils thus confirms that LSR signals in humans are not associated with the presence of memory B-cells (by contrast to CSR) and indicates that LSR is only transiently detectable contemporaneously to *in vivo* B-cell activation.

Human LSR junctions were related to either 3'RR1 or 3'RR2 and disseminated throughout these regions. The main conclusion to be drawn concerning this distribution is the high diversity of the breakpoints. The apparent higher frequency of breaks at the beginning of S μ likely relates to the preferential amplification of short fragments in a nested PCR assay. Breaks within the large 3'RR, also appeared to cluster in three sub-regions each spanning 3 to 5 kilobases, but these sub-regions in fact preceded the three unique non-repetitive locations where it had been possible to design highly specific reverse primers for the LSR-seq assay. While the use of additional regularly spaced reverse primers is prevented by the highly repetitive 3'RR sequence, it is thus likely that the frequency of breaks occurring at large distances from our primers is under-estimated.

Interestingly, sequence analysis of LSR-seq data from both *in vivo* and *in vitro* activated B-cells revealed the presence of both direct junctions from S μ to one or the other 3'RR, or of complex junctions including S γ or S α regions in addition to S μ .

Evaluation of LSR and comparison to CSR is inherently hindered by the need to identify junctions at the DNA level using nested PCR and sequencing (which can only be semi-quantitative), and by the fact that suiciding

cells rapidly vanish while switched cells are healthy and CSR remains detectable in the long term. In spite of that, we routinely identified LSR reads in DNA samples from activated B-cells, with frequencies up to 3% of the detection rate for CSR reads, meaning the LSR is a frequent event.

LSR junctions were found both in cells activated *in vitro* and in cells activated *in vivo* after CD40 ligation. It was not possible to identify conditions stimulating CSR in the absence of LSR. Even if variations appeared, with a tendency to increase the LSR/CSR ratio when TLR ligands were associated with CD40 ligation, or to decrease this ratio when BCR cross-linking was associated, these variations remained below statistical significance.

Another interesting point is the mechanisms of DNA breaks occurring in the 3'RR. While the process of LSR is globally AID-dependent, we cannot exclude that breaks in $S\mu$ might be initiated by AID while breaks in the 3'RR might be generated by another mechanism, yet to be identified. While the 3'RR includes repeats and is transcribed, similarly to a S region, our analysis showed that the distance between AID consensus motifs and 3'RR DSBs was much higher than for S regions. The same is true for the distance between breaks and G quadruplexes. This increase might in part relate to the structure of the target sequences: $S\mu$ is indeed extremely rich in AID hotspot motifs and G4 DNA; the mean distance of breaks to these target sites was accordingly very low (and did not differ significantly from the distance calculated under the hypothesis of random breaks). In contrast, both 3'RRs exhibit a lower density of AID hotspots and G4 DNA. 3'RR DSBs identified in unique junctions were located significantly closer to these sites than they would be if randomly distributed. This suggests active targeting of these sites by AID, even if the mean distance from AID sites is higher than for $S\mu$ breaks (or for $S\gamma$ breaks in CSR junctions). While the similarity of $S\mu$ breaks in CSR and LSR junctions strongly suggests that identical mechanisms are at work, the differences seen for the 3'RR might suggest that, in addition to the structure of the targeted sequence, either AID or its partners are differentially involved in both regions. The occurrence of breaks both within and outside LS repeats together with the lower density of AID sites suggest that, compared to $S\mu$, DSBs less easily follow staggered SSBs in the 3'RR and require more complex processing of initial SSBs. DNA cleavage outside of a repetitive target sequence has been previously reported for CSR to the $C\delta$ gene, where the target $\Sigma\delta$ sequence harbors few AID hotspots [31]. Even in classical S regions, once initiated in a repetitive region, R-loops eventually extend at distance though nearby G-rich regions, and hereby increases the length of the region targeted by CSR [25].

LSR was nearly absent in samples from AID-deficient individuals but we cannot exclude that this would only result from the absence of DSBs in $S\mu$. There is however no evidence indicating that enzymes other than AID might

contribute to LSR breaks, and notably the evaluation of distance from breaks to APOBEC3 family targets sites more or less ruled out that these deaminases might significantly contribute to LSR. Altogether, the significant implication of AID hotspots despite their relative scarcity implies that, in the 3'RR, cytosine deamination is followed by DNA remodeling on longer distances than within S_{μ} before occurrence of DSBs, synapsis to an S region and repair. The hypothesis of a change in the structure of the broken ends fits quite well with the observed increased length of microhomologies in repaired junctions. LSR is indeed repaired less frequently than CSR with blunt junctions typical of non-homologous end-joining, but with an increased usage of microhomologies and longer insertions, as described in the situation with increased involvement of alternate end-joining and microhomology mediated repair.

AID-initiated locus suicide recombination of the IgH locus thus actively occurs in human B-cells, with downstream breakpoints throughout the 3'RR super-enhancer. (Fig. 7). The phenomenon either predominates on the functional IgH allele or is bi-allelic, since LSR deletions are only detected transiently during B-cell activation but do not remain detectable in memory cells or plasma cells.

Because it is transient and followed by cell death, it remains difficult to evaluate LSR regulation and impact on B-cell homeostasis, relative to the other cell death pathways active in the GC. Few methods are available for studying deletions of hundreds kilobases with loosely determined breakpoints and the use of PCR followed with next generation sequencing remains semi-quantitative, given the PCR obstacles introduced by the repetitive nature of the 3'RR. In such settings, all conditions assayed for B-cell activation induce LSR at some degree and new methods will be necessary in order to determine where and when LSR occurs *in vivo* at the single-cell level.

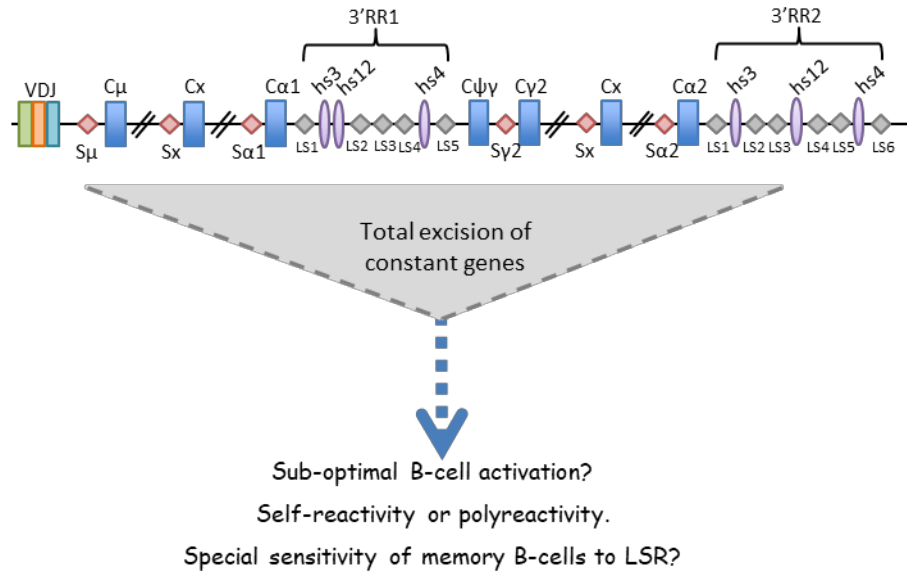


Figure 7: Model of LSR in the human IgH locus. Recombination between S_{μ} and LS regions in the 3'RR2 superenhancer leads to complete deletion of constant regions of all immunoglobulin classes leading to absence of BCR membrane expression. LSR could be a physiological mechanism to eliminate sub-optimally activated B-cells or those that are poly or autoreactive. Resting memory B-cells seem especially sensitive as they do not carry stigmas of LSR, which only occurs in activated B-cells.

Acknowledgements: We are grateful to M. Brousse, M. Faisse, S. Desforges and B. Remerand for technical help, and to Dr. Sven Kracker for help in the transfer of human samples. **Funding:** This work was supported by grants from ARC (PGA120150202338), Agence Nationale de la Recherche (ANR grant 16-CE15-0019-01), Institut National du Cancer (INCa grant #9363), Ligue Nationale contre le Cancer and Région Aquitaine-Limousin-Poitou-Charente. **Authorship Contributions:** MC designed the experiments, ID and ZD performed the experiments, AP, GL and TF provided B-cells from human lymphoid tissues and performed some in vitro activation assays. CD and FC provided RNAseq data from human activated B-cell samples. AD provided blood and lymphoid tissue samples from AID deficient individuals. FB performed *in silico* analysis of data and characterization of CSR/LSR junctions. JCM, SP, JCA and MC analyzed the data. JCM and MC wrote the manuscript. **Competing interests:** The authors declare that they have no conflict of interest.

REFERENCES

1. Mayer CT, Gazumyan A, Kara EE, Gitlin AD, Golijanin J, Viant C, et al. The microanatomic segregation of selection by apoptosis in the germinal center. *Science*. 2017;358. doi:10.1126/science.aao2602
2. Lebecque S, de Bouteiller O, Arpin C, Banchereau J, Liu YJ. Germinal center founder cells display propensity for apoptosis before onset of somatic mutation. *J Exp Med*. 1997;185: 563–571.
3. Goodnow CC, Vinuesa CG, Randall KL, Mackay F, Brink R. Control systems and decision making for antibody production. *Nat Immunol*. 2010;11: 681–688. doi:10.1038/ni.1900
4. Péron S, Laffleur B, Denis-Lagache N, Cook-Moreau J, Tinguely A, Delpy L, et al. AID-driven deletion causes immunoglobulin heavy chain locus suicide recombination in B cells. *Science*. 2012;336: 931–934. doi:10.1126/science.1218692
5. Pinaud E, Marquet M, Fiancette R, Péron S, Vincent-Fabert C, Denizot Y, et al. The IgH locus 3' regulatory region: pulling the strings from behind. *Adv Immunol*. 2011;110: 27–70. doi:10.1016/B978-0-12-387663-8.00002-8
6. Cogné M, Lansford R, Bottaro A, Zhang J, Gorman J, Young F, et al. A class switch control region at the 3' end of the immunoglobulin heavy chain locus. *Cell*. 1994;77: 737–747.
7. Pinaud E, Khamlichi AA, Le Morvan C, Drouet M, Nalesso V, Le Bert M, et al. Localization of the 3' IgH locus elements that effect long-distance regulation of class switch recombination. *Immunity*. 2001;15: 187–199.
8. Rouaud P, Vincent-Fabert C, Saintamand A, Fiancette R, Marquet M, Robert I, et al. The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. *J Exp Med*. 2013;210: 1501–1507. doi:10.1084/jem.20130072
9. Saintamand A, Rouaud P, Saad F, Rios G, Cogné M, Denizot Y. Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. *Nat Commun*. 2015;6: 7084. doi:10.1038/ncomms8084
10. Vincent-Fabert C, Fiancette R, Pinaud E, Truffinet V, Cogné N, Cogné M, et al. Genomic deletion of the whole IgH 3' regulatory region (hs3a, hs1,2, hs3b, and hs4) dramatically affects class switch recombination and Ig secretion to all isotypes. *Blood*. 2010;116: 1895–1898. doi:10.1182/blood-2010-01-264689
11. Le Noir S, Laffleur B, Carrion C, Garot A, Lecardeur S, Pinaud E, et al. The IgH locus 3' cis-regulatory super-enhancer co-opts AID for allelic transvection. *Oncotarget*. 2017; doi:10.18632/oncotarget.14585
12. Chauveau C, Cogné M. Palindromic structure of the IgH 3' locus control region. *Nat Genet*. 1996;14: 15–16. doi:10.1038/ng0996-15
13. Chauveau C, Pinaud E, Cogne M. Synergies between regulatory elements of the immunoglobulin heavy chain locus and its palindromic 3' locus control region. *Eur J Immunol*. 1998;28: 3048–3056. doi:10.1002/(SICI)1521-4141(199810)28:10<3048::AID-IMMU3048>3.0.CO;2-V
14. Saintamand A, Vincent-Fabert C, Garot A, Rouaud P, Oruc Z, Magnone V, et al. Deciphering the importance of the palindromic architecture of the immunoglobulin heavy-chain 3' regulatory region. *Nat Commun*. 2016;7: 10730. doi:10.1038/ncomms10730
15. Le Noir S, Boyer F, Lecardeur S, Brousse M, Oruc Z, Cook-Moreau J, et al. Functional anatomy of the immunoglobulin heavy chain 3' super-enhancer needs not only core enhancer elements but also their unique DNA context. *Nucleic Acids Res*. 2017;45: 5829–5837. doi:10.1093/nar/gkx203

16. Pinaud E, Aupetit C, Chauveau C, Cogné M. Identification of a homolog of the C alpha 3'/hs3 enhancer and of an allelic variant of the 3'IgH/hs1,2 enhancer downstream of the human immunoglobulin alpha 1 gene. *Eur J Immunol.* 1997;27: 2981–2985. doi:10.1002/eji.1830271134
17. Mills FC, Harindranath N, Mitchell M, Max EE. Enhancer complexes located downstream of both human immunoglobulin Calpha genes. *J Exp Med.* 1997;186: 845–858.
18. Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev.* 2004;18: 1618–1629. doi:10.1101/gad.1200804
19. Qiao Q, Wang L, Meng F-L, Hwang JK, Alt FW, Wu H. AID Recognizes Structured DNA for Class Switch Recombination. *Mol Cell.* 2017;67: 361-373.e4. doi:10.1016/j.molcel.2017.06.034
20. Caron G, Hussein M, Kulis M, Delaloy C, Chatonnet F, Pignarre A, et al. Cell-Cycle-Dependent Reconfiguration of the DNA Methylome during Terminal Differentiation of Human B Cells into Plasma Cells. *Cell Reports.* 2015;13: 1059–1071. doi:10.1016/j.celrep.2015.09.051
21. Caron G, Le Gallou S, Lamy T, Tarte K, Fest T. CXCR4 expression functionally discriminates centroblasts versus centrocytes within human germinal center B cells. *J Immunol.* 2009;182: 7595–7602. doi:10.4049/jimmunol.0804272
22. Boyer F, Boutouil H, Dalloul I, Dalloul Z, Cook-Moreau J, Aldigier J-C, et al. CSReport: A New Computational Tool Designed for Automatic Analysis of Class Switch Recombination Junctions Sequenced by High-Throughput Sequencing. *J Immunol.* 2017;198: 4148–4155. doi:10.4049/jimmunol.1601924
23. Bedrat A, Lacroix L, Mergny J-L. Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.* 2016;44: 1746–1759. doi:10.1093/nar/gkw006
24. Khamlichi AA, Glaudet F, Oruc Z, Denis V, Le Bert M, Cogné M. Immunoglobulin class-switch recombination in mice devoid of any S mu tandem repeat. *Blood.* 2004;103: 3828–3836. doi:10.1182/blood-2003-10-3470
25. Zhang ZZ, Pannunzio NR, Hsieh C-L, Yu K, Lieber MR. The role of G-density in switch region repeats for immunoglobulin class switch recombination. *Nucleic Acids Res.* 2014;42: 13186–13193. doi:10.1093/nar/gku1100
26. Vallur AC, Maizels N. Activities of human exonuclease I that promote cleavage of transcribed immunoglobulin switch regions. *Proc Natl Acad Sci USA.* 2008;105: 16508–16512. doi:10.1073/pnas.0805327105
27. Dalloul Z, Chenuet P, Dalloul I, Boyer F, Aldigier J-C, Laffleur B, et al. G-quadruplex DNA targeting alters class-switch recombination in B cells and attenuates allergic inflammation. *J Allergy Clin Immunol.* 2018; doi:10.1016/j.jaci.2018.06.011
28. Ribeiro de Almeida C, Dhir S, Dhir A, Moghaddam AE, Sattentau Q, Meinhart A, et al. RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Mol Cell.* 2018;70: 650-662.e8. doi:10.1016/j.molcel.2018.04.001
29. al GC et. Tissue distribution and dependence of responsiveness of human antigen-specific memory B cells. - PubMed - NCBI [Internet]. [cited 4 Aug 2018]. Available: <https://www.ncbi.nlm.nih.gov/insb.bib.cnrs.fr/pubmed/24567530>
30. Gadol N, Peacock MA, Ault KA. Antigenic phenotype and functional characterization of human tonsil B cells. *Blood.* 1988;71: 1048–1055.

31. Rouaud P, Saintamand A, Saad F, Carrion C, Lecardeur S, Cogné M, et al. Elucidation of the enigmatic IgD class-switch recombination via germline deletion of the IgH 3' regulatory region. *J Exp Med*. 2014;211: 975–985. doi:10.1084/jem.20131385

Article 2: Article soumis dans *The Journal of Experimental Medicine*.

Processes supported by the IgH 3'RR in mature B-cells, including hypermutation and recombination events, also rely on Mediator

Iman Dalloul, Sandrine Le Noir, Zeinab Dalloul, François Boyer, Bernardo Reina-San-Martin, Jeanne Moreau, Michel Cogné.

Dans les cellules B matures, le locus de la chaîne lourde des immunoglobulines (IgH) est contrôlé par des interactions dynamiques à longue distance qui régulent les événements de transcription et de recombinaison. Celles-ci sont soumises à une régulation stricte par un super-*enhancer* situé à l'extrémité 3' du locus (3'RR). La 3'RR exerce son rôle sur de longues distances afin de maintenir une activité élevée du promoteur VH et une expression optimale des immunoglobulines membranaires et sécrétées dans les cellules B matures et les plasmocytes, respectivement. Dans les lymphocytes B activés, la 3'RR contrôle également l'activation transcriptionnelle des régions constantes du locus et facilite le recrutement d'AID « Activation-Induced Cytidine Deaminase » à la région variable lors de l'hypermutation somatique (SHM), aux régions *switch* lors de la recombinaison de classe CSR « Class Switch Recombination » mais aussi sur elle-même, lors de la recombinaison suicide du locus LSR « Locus Suicide Recombination ».

Med1, une sous-unité du complexe Médiateur, s'est avérée faciliter l'activation transcriptionnelle du locus IgH mais aussi renforcer les interactions à longues distances qui ont lieu pendant le mécanisme de la CSR. Nous montrons ici que l'inactivation conditionnelle de Med1 dans les cellules B a un impact sur toutes les phénomènes contrôlés par la 3'RR. Nous montrons ainsi que Med1 favorise une expression optimale du BCR, il stimule la SHM et il est nécessaire pour l'hyper-transcription du locus IgH dans les plasmocytes. Med1 agit sur la CSR aussi bien lorsqu'elle survient en inter qu'en intra-chromosomique et ne semble pas avoir un rôle plus crucial pour des interactions plus « distantes », sans liaison physique covalente entre les 2 régions cibles situées sur des chromosomes différents.

L'absence de Med1 altère également fortement la transcription de la 3'RR et la LSR. Dans l'ensemble, nos résultats indiquent que le complexe Médiateur contribue à tous les effets

de la 3'RR qui nécessitent des interactions à longues distances et qui influencent l'état et le devenir du locus IgH dans les lymphocytes B matures.

Processes supported by the IgH 3'RR in mature B-cells, including hypermutation and recombination events, also rely on Mediator

Running title: Role of Mediator in *IgH* locus physiology

Iman Dalloul¹, Sandrine Le Noir¹, Zeinab Dalloul¹, François Boyer¹, Bernardo Reina-San-Martin^{2,3,4,5}, Jeanne Moreau¹, Michel Cogné,¹

¹Université de Limoges, CRIBL, UMR CNRS 7276, INSERM U1262, Limoges, France

² Institut de Génétique et de Biologie Moléculaire et Cellulaire; Illkirch, 67400; France.

³ Institut National de la Santé et de la Recherche Médicale, U1258; Illkirch, 67400; France.

⁴ Centre National de Recherche Scientifique, UMR7104; Illkirch, 67400; France.

⁵ Université de Strasbourg; Illkirch, 67400; France.

Keywords: Mature B-cells, Mediator, transcription, class switching, locus suicide recombination

Corresponding author: Michel Cogné, UMR CNRS 7276, University of Limoges, France.

Tel: 33 519564204; Fax: 33 555435896; email: cogne@unilim.fr;

ABSTRACT

In mature B-cells, the IgH locus is controlled by dynamic long-range 3D-interactions that regulate transcription and recombination events, under stringent regulation by the 3'RR super-enhancer. The 3'RR exerts its effects over long distances to maintain optimal expression of membrane Ig and secreted antibodies in B-lymphocytes and plasma cells, respectively. In activated cells, it controls recruitment of Activation-induced cytidine deaminase (AID) at three specific locations: the variable region during somatic hypermutation (SHM), switch regions during class switch recombination (CSR) and the 3'RR itself, during locus suicide recombination (LSR).

Med1, a subunit of Mediator, facilitates the transcriptional activation of the *IgH* locus and enforces the long-range interactions taking place during CSR. Here we show that, in addition, conditional inactivation of Med1 in B-cells impacts all of the features controlled by the 3'RR. Med1 promotes BCR expression, enhances SHM and is mandatory for hyper-transcription in plasma cells. We also show that Med1 promotes both inter- and trans-chromosomal CSR. Finally, the absence of Med1 appears to strongly impair 3'RR eRNA transcription and LSR. Overall, our results indicate that Mediator contributes to all of the long-range effects of the 3'RR that influence the status and fate of the *IgH* locus in mature B cells.

INTRODUCTION

The immunoglobulin heavy (*IgH*) chain locus encompasses more than two thousand kilobases and undergoes multiple remodelling, recombination and hypermutation events throughout differentiation and activation (Reynaud et al., 2008; Stavnezer et al., 2008; Dudley et al., 2005; Pinaud et al., 2011). During humoral immune responses, *IgH* expression varies profoundly in mature B-cells through various events occurring successively in germinal centers (GCs). Somatic hypermutation (SHM) of the rearranged variable regions begins and predominates in dark zone centroblasts, class switch recombination (CSR) of constant genes mostly marks light zone centrocytes, while cells engaged towards the plasma cell (PC) stage undergo focal hyper-transcription of *Ig* genes, thereby ensuring elevated antibody (Ab) production. In addition, activated B cells can also undergo a CSR-like form of recombination known as locus suicide recombination (LSR), which joins S_{μ} to the 3' *IgH* super-enhancer (3'RR) and which results in the loss of B cell receptor (BCR) expression and cell death (Péron et al., 2012),

These events are under the control of an *IgH* super-enhancer promoting not only transcriptional changes, but also chromatin remodelling, accessibility to recombinases, and recruitment of the SHM machinery (Cogné et al., 1994; Pinaud et al., 2011; Rouaud et al., 2013; Saintamand et al., 2015b). How these gene modifications remain focused on very specific targets within the locus still remains incompletely understood. By analogy to classical promoter-enhancer transcriptional interactions associated with looping of intervening sequences, the concept has emerged that intra-chromosomal loops are also needed prior to recombination. This is supported not only by the observation of strictly legitimate junctions between distant S-regions, their need for preexisting transcription, and by the direct demonstration of dynamic physical contacts between *IgH* enhancers, promoters and S-regions (Wuerffel et al., 2007; Thomas-Claudepierre et al., 2016).

While the concept of loops provides the simplest model representing the contacts between S regions, it overlooks the fact that junctions also occur between S regions from different *IgH* alleles, accounting for up to 10% of CSR events in mammals (Kingzette et al., 1998; Reynaud et al., 2005; Laffleur et al., 2014). Accommodating all observed events thus needs to consider that the general prerequisite to CSR is the occurrence of legitimate interactions between paired S regions, whether in *trans* or in *cis* and only the latter result in DNA looping and fit with the prevalent model. Coexistence of these classical *cis*-CSR and alternate *trans*-CSR pathways supports the concept of “CSR factories”, as areas attracting any DNA sequence targeted, bound, or functionally related to the CSR machinery, whether or not bounding an intra-chromosomal loop, and regardless of being an S-region or the *IgH* 3' super-enhancer. Since *trans*-interactions should not yield chromosomal instability, a major issue is to ensure their legitimacy: interallelic *IgH* interactions preceding CSR, and resulting allelic transvection, might contribute to CSR legitimacy (Le Noir et al., 2017). It is currently unknown whether *trans*-interactions are looser than loops occurring in *cis* and differentially regulated by specific factors. While LSR can be considered as a variation of the CSR process, the parameters governing the choice between CSR and LSR are unknown and detailed study of the conditions regulating the synapsis of AID targets over long distances is necessary.

Mediator is a large (four-modules) multi-subunit complex with an important role in connecting DNA-bound transcription factors directly to RNA polymerase II (Pol II). It controls Pol II C-terminal domain (CTD) phosphorylation and communicates various regulatory signals to the polymerase, in forming the preinitiation complex, pausing, or elongation (Allen and Taatjes, 2015). It also interacts with the Cohesin complex, itself known to stabilize chromosome pairing at meiosis, and was thus suggested to promote interactions between chromosomes of the same pair (Carlsten et al., 2013). Notably, the Med1 subunit interacts with various transcriptional activators, the recruitment of which to chromatin culminates at locations of super-enhancers

and marks them with a one-order-of-magnitude higher load of Mediator than for simple enhancers (Whyte et al., 2013). As for any Mediator subunit specific for a given set of transcription factors, Med1 is only included in a fraction of all Mediator complexes and it is not mandatory for transcription (Allen and Taatjes, 2015). Its activity additionally relies on its phosphorylation by the MapK/ERK pathway, followed by increased recruitment to Mediator complexes and boosted transcription of specific genes in activated cells (Belakavadi et al., 2008). Interestingly, Med1 was recently shown to qualitatively affect SHM and to promote germline transcription of switch (S) regions and long-range enhancer/S region interactions during CSR (Thomas-Claudepierre et al., 2016; Gaudot et al., 2018). Mediator thus stands as a good candidate for connecting the MAPK/ERK-dependent B-cell activation and all the subsequent *IgH* locus remodelling and 3D changes controlled by the 3'RR super-enhancer during late B-cell differentiation. Given the functional importance of gene architecture and long-range interactions for the physiology of the *IgH* locus, we decided to explore the role of the Med1 subunit of Mediator on all the features relying on the 3'RR in mature activated B cells and plasma cells: BCR expression, SHM, CSR (*cis* and *trans*), LSR and antibody secretion.

Results

Decreased membrane BCR expression and antibody secretion in Med1-deficient mature B-lineage cells.

Conditional ablation of Med1 in the B-cell lineage is known to globally preserve B-cell development (Thomas-Claudepierre et al., 2016). While all splenic B-cell populations were present, we observed a significantly lower density of membrane BCR in almost all B-cell compartments, and most significantly in IgM^{high} marginal zone B-cells and B1 cells (**Fig. 1A**). Consistent with this, we found that the serum levels of circulating antibodies was lower in Med1-deficient mice (**Fig. 1B**). This was particularly noticeable for IgA and IgE, short-lived Ig not recycled by the FcRn pathway, which were cut by up to five-fold (**Fig 1B**). Although less important, a significant (about 2-fold) reduction in serum IgG and IgM also manifested. Furthermore, single-cell analysis of antibody secretion by Elispot showed that bone marrow-derived plasma cells isolated from Med1-deficient mice showed a significant decrease in the number of IgM-expressing spots (approximately cut by half) and area (cut by one-third) when compared to wild type controls (**Fig1C, Top**). As expected from the CSR defect described for Med1-deficiency (Thomas-Claudepierre et al., 2016), IgG1 plasma cell numbers were decreased to a much higher extent in Med1-deficient mice, and again showed reduced spot areas, indicating decreased levels of secretion on a per cell basis (**Fig1C, bottom**). We conclude that Med1-deficiency impacts the level of BCR expression and antibody secretion in mature B cells and plasma cells.

Med1-deficiency results in decreased levels of both intra- and inter-chromosomal CSR and severely impaired LSR.

In agreement with the previously described CSR defect after *in vitro* stimulation of Med1-deficient B cells (Thomas-Claudepierre et al., 2016), we found that LPS/IL4 stimulated B-cells

secreted less IgG1 in supernatants, yielded a lower number of IgG secreting cells by Elispot (while IgM producing cells appeared correlatively higher) and globally showed lower transcription of C γ 1 (**Fig. S1**). To gain insight into the CSR defect at the molecular level and further appreciate how much Med1-deficiency impacts on CSR and to also determine whether it affects LSR, we sequenced CSR and LSR junctions detectable in LPS/IL4 activated B cells using a semi-quantitative nested PCR assay coupled to high-throughput sequencing (Boyer et al., 2017). We used primers flanking the S μ and S γ 1 switch regions to amplify CSR junctions and primers flanking S μ and the 3'RR to amplify LSR junctions (adapted from our initial description of LSR (Péron et al, 2012)). As expected, Med1-deficient B cells yielded a significantly lower level of CSR junctions when compared to controls (**Fig. 2A, left**). This was true when considering either the number of unique reads identified as CSR junctions, or the numbers of “unique junctions” (repeatedly found in multiple reads).

Surprisingly, we found that the effect of Med1-deletion on LSR is much more pronounced than on CSR, as LSR junctions were barely detectable in Med1-deficient B-cells (**Fig. 2A, right**).

The repair of CSR breaks is known to influence the fine structure of CSR junctional sequences. Those sequences obtained from Med1-deficient B-cells significantly varied from *wt* controls by showing an increased frequency of short junctional insertions together with increased usage of microhomology-mediated repair (with consequently a relative decrease of blunt junctions compared to *wt* controls) (**Fig. 2B**).

We conclude that Med1 deficiency quantitatively impairs the efficiency of CSR and LSR, and also qualitatively impacts the repair of CSR junctions.

In order to evaluate physical interactions between *IgH* alleles, inter-allelic distances were measured by 3D-FISH and compared in resting *versus* activated B-cells in either wild-type or Med1-deficient B-cells. Short-range inter-allelic interactions were, as expected, increased upon

B-cell activation, but they did not significantly differ in Med1-deficient compared to -proficient cells, neither in LPS-stimulated nor in unstimulated B-cells (**Fig. 2C**).

IgH alleles are known to frequently undergo reciprocal exchanges during CSR. To determine whether Med1-deficiency might more severely impact *trans* CSR and junctions between different chromosomes, we used mice heterozygous for the *IgH* a/b allotypes and measured the frequency of switched VDJ-C α *IgH* transcripts showing the association of a J_H^a allotype with C α ^b allotype (or reciprocally of J_H^b with C α ^a) in B cells purified from the Peyer's patches. This allowed us to evaluate the proportions of *trans*-CSR vs. *cis*-CSR for the C α gene. We found that the proportion of *trans*-CSR and *cis*-CSR was not altered by Med1-deficiency (**Fig 2D**). Therefore, Med1-deficiency equally affects recombination in *cis* and in *trans*.

Med1-deficiency affects the frequency of somatic hypermutation in coding regions.

To assess whether Med1-deficiency also functionally affects SHM in coding gene segments, we sequenced expressed VDJ regions from Peyer's patch B-cells. Sequences were specifically analyzed for SHM in class-switched *IgH* transcripts. We found a two-fold decrease in the frequency of SHM, affecting both the framework and complementary determining regions of VDJ rearranged exons (**Fig. 3A**). Globally 52437 mutations (37455 replacements (R) and 14982 silent (S) mutations) were found among 2231 kb of sequenced V regions from *WT* controls (23.5 per Kb), while only 46823 mutations (33445 R and 13378 S mutations) were found among 3337 kb of sequenced V regions from Med1-deficient B-cells (14.0 per Kb). By contrast, the AT/GC mutation pattern in coding regions was not different between sequences obtained from Med1-deficient B cells and controls when considering all mutations or when analyzing R or S mutations separately (**Fig. 3B**). We conclude that Med1-deficiency affects the frequency of SHM in coding regions without altering the mutation pattern.

Transcription of the IgH 3'RR is impaired in the absence of Med1.

As for other enhancers, accessibility and activity of the *IgH* 3'RR can be assessed by evaluating transcription of core enhancer elements into eRNA. Comparison of Med1-proficient and -deficient B-cells for *IgH* transcripts after a 4-days stimulation by LPS revealed a major reduction in eRNA transcripts from the 3'RR hs1,2 and hs4 core enhancer elements, which were reduced by three and five-fold, respectively. This was more pronounced than for the primary *IgH* transcripts, which were reduced by two-fold in parallel evaluation of transcription in the J_H-E_μ intron (**Fig. 4A**).

Chromatin accessibility throughout the *IgH* locus, including the E_μ and 3'RR enhancers, was assessed by chromatin immunoprecipitation (ChIP) experiments on LPS-activated B cells and using antibodies specific for histone marks classically associated with active enhancers (H3K9 and H3K27 acetylation). By contrast to 3'RR transcriptional changes mentioned above, we found no difference in H3K9 and H3K27 acetylation between Med1-deficient and -proficient B cells (**Fig 4B**). The only apparent change was a higher signal of H3K27 acetylation at the S_μ region, most likely related to the more frequent preservation of S_μ in LPS-activated Med1-deficient cells due to the CSR defect.

Discussion

The Med1 subunit of Mediator is involved in functional and physical loops between promoter and enhancer elements. Besides its role in promoting transcription within chromosomal loops, it is also postulated to play a role in inter-chromosomal interactions, notably after its abundant binding to super-enhancers which broadly influences 3-D nuclear architecture (Whyte et al., 2013). Med1 was recently shown to partially impact long-distance interactions within the *IgH* locus during CSR (Thomas-Claudepierre et al., 2016). A mild effect on SHM was also found in the intron following a rearranged JH element, without global SHM decrease but with a biased A-T / G-C ratio in Med1-deficient B-cells (Gaudot et al., 2018). Not only CSR and SHM but also multiple quantitative and qualitative traits of *IgH* expression are under control by chromosomal loops coordinated by the *IgH* 3'RR super-enhancer (Wuerffel et al., 2007). We thus explored in detail the impact of Med1 on both the amount and the quality of expressed Ig in mature B-cells and on the various *IgH* locus features controlled by long-distance interactions with the 3'RR super-enhancer.

This revealed that the phenotype of Med1-deficiency is reminiscent of the phenotype of 3'RR-deficient mice (Saintamand et al., 2015a; Ghazzoui et al., 2018; Rouaud et al., 2013; Vincent-Fabert et al., 2010), although milder. In addition to the previously reported partial class-switch defect (Thomas-Claudepierre et al., 2016), Med1-deficient B-cells also featured a decreased density of BCR expression affecting all B-cell compartments and most notable in B1 and MZ cells. This is very similar to the decreased BCR density observed in 3'RR-deficient animals, which does not affect signaling and B-cell activation but associates with slightly reduced basal transcription of the pVH promoter (Saintamand et al., 2015a). Regarding Ig secretion, which normally correlates with strongly boosted Ig gene transcription at the plasma cell stage, we also observed that the absence of Med1 led to defective secretion of all Ig classes *in vivo*, affecting

not only class-switched Ig (IgG, IgA, IgE) but also, to a lesser extent, IgM. The decrease in serum IgG is especially significant since IgG has a long half-life *in vivo* and usually accumulates even in conditions of lower production due to homeostatic regulation of recycling and catabolism (Fahey and Sell, 1965). Again, the globally decreased Ig secretion including IgM phenocopies the alterations seen in 3'RR-deficient mice, since the 3'RR super-enhancer is mandatory for full-blown *IgH* expression in plasma cells (Vincent-Fabert et al., 2010).

Besides quantitative aspects, optimal immune responses also rely on qualitative changes triggered by antigen encounter, such as Ig affinity maturation and associated hypermutation of V region sequences. In activated mature B-cells, SHM of *IgH* V regions stands as another process supported by the *IgH* super-enhancer (Rouaud et al., 2013). In this regard, a noticeable difference in SHM appeared between expressed *IgH* variable regions from either Med1-deficient or proficient mice. Med1-deficient conditions revealed a quantitative SHM defect affecting all segments of the VDJ region, both framework and complementarity determining regions. This defect most likely resulted from mechanistic alterations of SHM and was not at the cell selection level, since it similarly affected silent and replacements mutation. The evaluated functional sequences from exons, were by contrast not affected by the bias of AT vs GC mutation that we previously noticed in the J_{H4} flanking intron, which might be restricted to non-coding regions (Gaudot et al., 2018).

Another more functional aspect of inter-allelic *IgH trans*-interactions is the occurrence of *trans*-recombination during CSR. Indeed, while intrachromosomal repair predominates in B-cells in the case of IsceI-induced breaks (Gostissa et al., 2014), physiological CSR involves a significant ratio of *trans*-junctions between allelic chromosomes. *Trans*-CSR, a legitimate and functional process, is then much more frequent than rare junctions with illegitimate partners (such aberrations being connected to lymphomagenesis). We checked whether *trans*-CSR was more dependent upon Med1 than *cis*-CSR by sequencing α heavy chain transcripts and

quantifying the re-pairing of J_H and C α allotypes in *IgH* a/b heterozygous mice. *Cis* and *trans*-CSR were equally affected in Med1-deficient mice compared to *wt*. Although Mediator binds both the Cohesin complex and super-enhancers suspected to organize nuclear 3D-architecture, Med1 thus does not appear to be mandatory neither for interallelic pairing nor for legitimate inter-chromosomal *trans*-CSR junctions.

Besides such internal variations in the mechanistics of functional CSR, a last and dramatic AID-dependent recombination of *IgH* constant genes has also been described. LSR, which deletes constant genes in some activated B-cells thus stands as an alternative to CSR with the potential to delete the relevant activated B-cells (Péron et al., 2012). Evaluation of this process by LSR-seq, a high-throughput analysis of recombined DNA fragments, showed a drastically reduced amount of LSR junctions, still more pronounced than for CSR junctions.

Interestingly the sequences of CSR junctions revealed an additional anomaly, with increased frequency of microhomology and of junctional insertions, both hallmarks of alternative end-joining (A-EJ). This suggests that beside the predominant non-homologous end-joining (NHEJ) pathway Med1-deficient B-cells make increased usage of A-EJ during the repair step of CS,.

Except for this bias in the repair process, a common feature of all other the above-mentioned processes altered in Med1-deficient cells is their dependence upon 3'RR super-enhancer function, either in basal conditions where the 3'RR optimizes *IgH* transcription and BCR expression in mature B-cells, or in the conditions of B-cell activation by antigen or TLR ligands. When checking the activity of the 3'RR directly, we observed that a known marker of enhancer activity, eRNA transcription (Wang et al., 2011), was significantly decreased. This is consistent with decreased functional effects of the super-enhancer and with its decreased availability as a recombination target during LSR. By contrast, the major chromatin marks associated with enhancer function, H3K9 and H3K27 acetylation, were unchanged, indicating that the super-

enhancer was normally bound by the transcription factors recruiting histone acetylation transferases. In these experiments, eRNA transcription of core enhancer elements was the most notable change and Med1-deficiency thus uncoupled the various parameters related to a super-enhancer active status. Since Med1 function strongly depends upon its phosphorylation by ERK in activated cells, and since 3'RR eRNA are at the highest levels in activated B-cells, these observations suggest successive steps of 3'RR activation, with an initial and partial activation status associated with chromatin remodeling, followed by induction of Med1-dependent *hIgH* eRNA transcription in activated B-cells. Only a super-enhancer optimally loaded with Pol II following interaction with a distant promoter and Mediator, might fully support 3'RR functions. The 3'RR super-enhancer is a large and complex DNA region indeed carrying multiple functions, either simultaneously (germline transcription and CSR), successively (hypertranscription in PCs) or alternatively (SHM *vs* CSR *vs* LSR...), thus requiring the identification of conditions differentially affecting or uncoupling its functions.

It is noticeable that SHM and CSR are reduced but not abrogated in the absence of Med1, indicating that functional interactions of 3'RR with *IgH* V and S regions are maintained, in agreement with previous 4C-Seq data showing persistent E μ -3'RR contacts in the absence of Med1 (Thomas-Claudepierre et al., 2016). Med1 thus extends or enforces chromosomal looping and long-range interactions of target regions with the relevant super-enhancer, without being completely mandatory.

LSR, featuring junctions of double strand breaks in the 3'RR to those within S μ , appears as the process most affected by a Med1 defect, *i.e.* the most dependent upon Mediator. This suggests an implication of Med1 in the alternate choice between CSR and LSR. The parallel strong reduction in 3'RR eRNA might indicate that either Med1 itself and/or the *hIgH* level eRNA transcription resulting from Med1-Pol II interactions are prerequisites for LSR.

While *IgH* locus physiology largely relies on 3D long-range interactions, our results further demonstrate that beyond a general role on transcription, Mediator participates in the regulation of all processes governed by the *IgH* 3'RR super-enhancer, including SHM, super-transcription in plasma cells, CSR (*cis* & *trans*) and LSR.

Material and methods

Mice

Med1 fl/fl mb1-cre mice have been previously reported (Thomas-Claudepierre et al., 2016) and were obtained from Institut de Génétique et de Biologie Moléculaire et Cellulaire, Strasbourg, France). Control *WT* C57B16/SV129 mice were purchased from Charles River. Procedures reviewed and approved by the local ethical review board.

Blood sampling

Blood samples were recovered from ten-week old mice with heparinised needles .Plasma samples were recovered by centrifugation and stored at -20°C until used.

Cell preparation

Splenocytes were collected, red blood cells were lysed and B-cells were isolated using EasySep™ Mouse B-cell Isolation Kit (Stem cell). B-cells were cultured for 2 days (for Chip experiments), 3 days (for 3D-FISH) and 4 days (for ELISA, Elispot and qPCR), in RPMI containing 10% fetal bovine serum with LPS (1 µg/mL) (Invivogen) or LPS + IL-4 (20 ng/mL) (Peprotech). Supernatants were recovered and stored at -20°C until used. Cultured splenic B-cells were used in parallel for RNA preparation.

Plasma cells were enriched from bone marrow using CD138 microbeads (Miltenyi biotech).

Class-specific ELISA

ELISA was performed on sera and supernatants from *in vitro* stimulated primary B-cells for the detection of the various Ig classes and subclasses. Plates were coated overnight with monoclonal antibodies specific for IgM, IgA, IgE, IgG1, and total IgG (Southern Biotech). Sera or supernatants were added and incubated for 2h at 37°C. After washing, alkaline phosphatase

(AP)-conjugates of goat anti mouse IgM, IgG1, IgE, IgA and total IgG (Southern Biotech) were incubated for 1h at 37°C. Following washing and addition of AP substrate (Sigma), absorbance was measured at 405 nm.

Elispot

For evaluation of IgG1 or IgM secretion, cells (*in vitro* activated B-cells or enriched plasma cells from bone marrow) were seeded in triplicate in culture medium in a 96-well MultiScreen HTS plate (Millipore) coated with 100ng/well of anti-IgG1 or anti-IgM. Cells were incubated overnight at 37 °C and then removed by washing with PBS/0.01 % Tween. Plates were then incubated 1 h at 37 °C with AP-coupled anti-IgG1 or phosphatase-coupled -anti-IgM (1/1000). After washing, addition of 100 µL BCIP/NBT AP substrate (Millipore), new washes and drying, images were taken with an NI-E microscope (Nikon) for evaluation of spot numbers and areas with the Nis-Ar software (Nikon).

Flow cytometry

Single-cell suspensions from fresh spleens were washed with PBS containing 2% FCS. Mature B-lymphocytes were stained with CD19-APCH7, CD5-FITC, CD21-PE, CD23-PECy7 (BD Pharmingen), IgD-BV421 (Biolegend) and IgM-APC (e Bioscience) antibodies. *IgH* allotype expression on peripheral blood B-cells was verified using antibodies directed against IgM^a (IgM^a-FITC: BD Pharmingen) and IgM^b (IgM^b-PE: BD Pharmingen) allotypes. Flow cytometry was performed on a BD LSR Fortessa cell analyzer (BD Biosciences). FACS data were analyzed using FACSDIVA software (BD Biosciences).

Real time quantitative PCR (RT-qPCR)

Quantitative PCR was performed using SYBR[®] qPCR green and premix Ex Taq (TAKARA). Hs1.2, hs4, IgH, post-switch IgG1, membrane and secreted IgM transcripts were quantified after four days of B-cell stimulation with LPS or LPS+IL4. RNA from B-lymphocytes was extracted using GenElute Mammalian RNA miniprep Kit (Sigma Aldrich). cDNA was synthesized using the High Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific). Quantitative PCR primers used to study heavy chain expression for the secretory IgM (mIgM qPCR For, mIgM qPCR SEC Rev), membrane IgM (mIgM qPCR for, mIgM qPCR MB Rev), IgG1 (I μ for primer, C γ 1 Rev primer) and rRNA transcription from hs1,2 (hs1,2 For, Hs12 Rev) and hs4 (hs4 For, hs4 Rev) are listed in Supplemental Table1. *IgH* primary transcripts were quantified using Taqman probes as previously described (Rouaud et al, JEM 2013). All transcripts were normalized to Glyceraldehyde-3-phosphate dehydrogenase (Gapdh) transcripts (reference Mm99999915-g1).

ChIP experiments

ChIP experiments were done on LPS stimulated sorted CD43^{neg} spleen B-cells. Briefly, 15 \times 10⁶ cells were cross-linked at room temperature for 15 min in 15 mL PBS, 1% formaldehyde. The reaction was quenched with 2.125M glycine. After lysis, chromatin was sonicated to 0.5–1 kb using a Vibracell 75043 (Thermo Fisher 428 Scientific). Following 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 2h at 4 °C with 50 mL of 50% protein A/G slurry (0.2 mg / mL sheared salmon sperm DNA, 0.5 mg/ mL BSA, and 50% protein A/G; Sigma). Cell equivalents (1 \times 10⁶) were saved as input, and the remaining cell equivalents were incubated overnight with anti-acetyl-histone H3(Lys9) (Merck) or with anti-acetyl-histone H3(Lys27) (Abcam) polyclonal antibodies. Immune complexes were precipitated by the addition of protein A/G. Cross-linking was reversed by overnight incubation (70 °C) in TE

buffer with 0.02% SDS and chromatin was phenol/chloroform extracted. QPCR assays used for evaluating precipitated DNA from E μ , S μ , S γ 3, hs3a, hs3b, hs12 and hs4 are listed in Supplemental Table 1.

Amplification of S μ /S γ and S μ /LS4 junctions by NGS

DNA from LPS+IL-4 stimulated cells (from *WT* and *Med1mb1Cre* mice) was extracted using GenElute Mammalian Genomic DNA miniprep Kit (Sigma Aldrich). S μ /S γ and S μ /LS4 junctions were amplified in triplicate by nested PCR with 100 ng DNA (Phusion HF polymerase; BioLabs) using the following primers: S μ Nest1 For and S γ consensus1 Rev for PCR1 S μ /S γ junctions and S μ Nest2 For and S γ consensus2 Rev for PCR2. S μ /LS4 junctions were amplified using Smum1 For and LS4-1 Rev primers for PCR1 and Smum2 For and LS4-2 Rev for PCR2. Primers used for amplification of these junctions are listed in Supplemental Table 1. Each library was prepared using 200 ng PCR2 product. Barcoded libraries with 200-pb read lengths were prepared using Ion Xpress plus Fragment Library Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Each barcoded library was mixed in equal amounts and diluted to 100 pM. Libraries were run on an Ion PI v3 chip on the Ion Proton sequencer (Life Technologies). Data were analyzed using the CSReport software (Boyer et al., 2017).

Illumina sequencing for somatic hypermutation SHM and trans-switch

SHM and *trans*-switch analyses were performed on Peyer's patches (IgH^a/IgH^b allotype). RNA was extracted using GenElute Mammalian RNA miniprep Kit (Sigma Aldrich). These experiments combined RACE PCR and library preparations. RACE PCR was performed on 500ng RNA incubated at 72°C for 3 min and at 42°C for 2 min with a specific IgA reverse primer and 1 μ l dNTP (10Mm) (MP Biomedicals). Caprace forward primer, DTT, 5X ProtoScript buffer and ProtoScript enzyme (BioLabs) were added to the mixture and incubated at 42°C for 90 min and then at 70°C for 10 min. Libraries were prepared on RACE products

using two forward primers: Universal short For and Universal long Forward primers and one IgA reverse specific primer (Mm-C α +link). Primer sequences used in this experiment are reported in Supplemental Table 1. Amplification was performed with Phusion[®] High-Fidelity DNA Polymerase (New England Biolabs) according to the following program: DNA was denatured 30 s at 98°C and then submitted to 30 cycles consisting of 98°C for 10 s, 65°C for 30 s and 72°C for 30 s, and 1 cycle at 72°C for 10 min. Adaptors were added and the resulting amplicons were sequenced using an Illumina Miseq sequencer. For analysis, the international ImMunoGeneTics information system (IMGT), IMGT/HighV-QUEST Web portal, and IMGT-ONTOLOGY concepts were used.

In order to identify alleles in the *JH1* and C α CH1 regions, raw reads R2 from Illumina paired-end sequencing were first filtered by their size (min. 150 bp) and aligned against Ighj-IgHa rearranged sequences (last 34 bp of Ighj + first 76 bp of IgHa) using *pairwise2* local alignment function of Biopython package (implementation of Smith-Waterman algorithm). Reference sequences were composed of either Igh1*a, Igh1*b, Igh2, Igh3, or Igh4 and Ighj gene/allele was assigned to each read according to best alignment score. Identification of *IgH* α chain alleles was based on the aligned base at position 43 in IgHa (C/G polymorphism between Igha*a/Igha*b alleles). Then *cis* junctions were counted as the number of occurrences of either Ighj1*a-IgHa*a (a-a) or Ighj1*b-IgHa*b (b-b), and *trans* junctions as the number of occurrences of either Ighj1*a-IgHa*b (a-b) or Ighj1*b-IgHa*a (b-a).

3D-FISH

In vitro activated B-cells were dropped onto poly-L-lysine slides and fixed with 4% paraformaldehyde for 10 min at room temperature (RT). After washing with PBS, cells were permeabilized with 0.02% pepsin / 0.1M HCl for 15 min at RT, then washed and post-fixed with 1% paraformaldehyde for 5 min at RT, denatured for 5 min in 70% formamide, 2X SSC

at 72°C and hybridized overnight at 37°C. The *IgH* (RP23-109B20) probe (encompassing an AID-targeted region from D_H cluster to C_γ2a) was labelled with dCTP-biotin (Invitrogen). Slides were washed in 1X SSC at 72°C for 5 min, incubated with streptavidin-Alexa Fluor 594 (1/200, BD biosciences) 1h at RT, and mounted with Vectashield (Vector labs). Images were acquired with an epifluorescence microscope (NIKON). Optical sections separated by 0.2µm were collected and stacks were de-convoluted and analyzed using Huygens and Volocity software, respectively. Volumetric pixel size was 0.064µm in xy and 0.2 µm in z.

Statistical analysis:

Statistical tests were performed using GraphPad Prism (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001)

Conflict of interest – None

Acknowledgments - This work was supported by grants from Conseil Régional du Limousin, Association pour la Recherche sur le Cancer (PGA120150202338), ANR Ig-MemImpact (16-CE15-0019-01), ANR IgH-AID-access and ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the frame program Investissements d’Avenir ANR-10-IDEX-0002-02. F. Boyer was supported by a grant from ALURAD. We thank S. Desforges and B. Remerand for help with animal care.

Contribution section - M. Cogné, J Cook-Moreau actively participated in the experimental design of the study. I. Dalloul, S. Le Noir and Z. Dalloul performed experiments. F. Boyer analysed Ig repertoire data. B. Reina-San-Martin provided Med1-deficient mice and participated in the scientific discussion and manuscript writing.

FIGURE LEGENDS

Figure 1: Membrane IgM expression and Ig secretion *in vivo* in Med1^{F/F}Mb1^{Cre/+} mice. (A) Peripheral B-cell populations were monitored for IgM BCR expression on *wt* and Med1^{F/F}Mb1^{Cre/+} mice by flow cytometry measuring mean fluorescent intensity. Data represent mean ± SEM from 6 *wt* and 6 Med1^{F/F}Mb1^{Cre/+} mice. (B) Serum IgM, G, A and E concentrations were evaluated by ELISA in ten-week-old mice. Data represent mean ± SEM from at least 5 mice. (C) Elispot specific for IgM and IgG1 secretion, mean number of spots (left) and mean area (right) of spots were measured for bone marrow plasma cells from *wt* and Med1^{F/F}Mb1^{Cre/+} mice. Data are from 2 independent experiments with at least 3 control and 3 Med1^{F/F}Mb1^{Cre/+} mice in each experiment. Mann-Whitney test was used for significance.

Figure 2: CSR and LSR in Med1^{F/F}Mb1^{Cre/+} mice. (A) S μ -S γ 1 CSR junctions and LSR junctions were PCR-amplified and sequenced from LPS+IL4 stimulated spleen B-cells, comparing mutant and control mice. Data are mean counts ± SEM of identified junctions. (B) Junction structures were analyzed using CSReport. Data are means obtained from groups of 5 mice. (C) Distribution of *IgH* inter-allelic distance (bottom) was obtained by 3D-FISH in spleen B-cells (Day 0) and LPS stimulated B-cells (Day3) (with normalization according to cell diameter) in control and Med1^{F/F}Mb1^{Cre/+} mice. Data represent mean distances ± SEM calculated from 2 independent experiments, each with 3 control and 3 Med1^{F/F}Mb1^{Cre/+} mice. KS test was used for significance. (D) *Cis* and *trans*-switch analysis was performed on Peyer's patch RNA (from heterozygous IgH^a/IgH^b mice) using a specific C α CH1 reverse primer. The various *cis* and *trans* associations carrying IgH^a/IgH^a, IgH^b/^b (*cis*) and IgH^a/^b (*trans*) were counted after NGS sequencing. Data represent means ± SEM from 7 control and 11 Med1^{F/F}Mb1^{Cre/+} mice. Mann-Whitney test was used for significance. Chi2 test was used for significance of IgG1 CSR structure analysis.

Figure 3: SHM in Med1^{F/F}Mb1^{Cre/+} mice. (A) SHM of VDJ regions was studied in Peyer's patch B-cells after NGS sequencing using a C α specific reverse primer. 7746 clonotypes in control mice (n=7) and 11724 clonotypes for Med1^{F/F}Mb1^{Cre/+} mice (n=11) were characterized. Mutation rates per 1kb in the full V region and in the various subdomains (FR1, CDR1, FR2, CDR2 and FR3) are shown (A), together with mutation spectra (B) either shown globally (left) or separately for R (middle) and S mutations (right), comparing *wt* controls and Med1^{F/F}Mb1^{Cre/+} mice. Mann-Whitney test was used when comparing mutation rates; Chi2 test was used for mutation spectra comparison.

Figure 4: *IgH* locus transcription and epigenetic marks in Med1^{F/F}Mb1^{Cre/+} mice. (A) Primary IgH, *hs1,2* and *hs4* transcripts were evaluated by RT-qPCR of RNA from splenocytes stimulated with LPS for 3 days. (B) ChIP assays with anti-H3K9ac and anti-H3K27ac Ab were performed in 2-days LPS-stimulated B-cells from control and Med1^{F/F}Mb1^{Cre/+} mice. (C) Recruitment of H3K9ac (top) and H3K27ac (bottom) to the *IgH* locus at E μ , S μ , S γ 3, *h3a*, *h1-2*, *hs3b* and *hs4* regions was analyzed by QPCR. Data represent mean signal ± SEM from 6 mice. Mann-Whitney test was used for significance.

Supplementary figure 1: Ig secretion by Med1^{F/F}Mb1^{Cre/+} *in vitro* stimulated splenocytes. (A) B-cells were stimulated for 4 days with LPS and IL-4. Supernatants were quantified by ELISA for IgM, IgG1 and IgE. Data represent mean concentrations ± SEM from 3 independent experiments, each with at least 3 control and 3 Med1^{F/F}Mb1^{Cre/+} mice. (B) Elispot specific for

IgM (left) and IgG1 (right) secretion was performed on LPS+IL4 stimulated B-cells from control and Med1^{F/F}Mb1^{Cre/+} mice. Data represent mean numbers \pm SEM from 2 independent experiments, each with at least 3 control and 3 Med1^{F/F}Mb1^{Cre/+} mice. **(C)** Transcripts encoding membrane (mb) and secreted (sec)-form of μ and γ 1 chains were quantified by RT-qPCR in B-cells stimulated for 4 days with LPS+IL-4. Data represent relative expression \pm SEM from 3 control and 4 Med1^{F/F}Mb1^{Cre/+} mice. Mann-Whitney test was used for significance.

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. doi:10.1038/nrm3951.
- Belakavadi, M., P.K. Pandey, R. Vijayvargia, and J.D. Fondell. 2008. MED1 phosphorylation promotes its association with mediator: implications for nuclear receptor signaling. *Mol. Cell Biol.* 28:3932–3942. doi:10.1128/MCB.02191-07.
- Boyer, F., H. Boutouil, I. Dalloul, Z. Dalloul, J. Cook-Moreau, J.-C. Aldigier, C. Carrion, B. Herve, E. Scaon, M. Cogné, and S. Péron. 2017. CSReport: A New Computational Tool Designed for Automatic Analysis of Class Switch Recombination Junctions Sequenced by High-Throughput Sequencing. *J. Immunol.* 198:4148–4155. doi:10.4049/jimmunol.1601924.
- Carlsten, J.O.P., X. Zhu, and C.M. Gustafsson. 2013. The multitasked Mediator complex. *Trends Biochem. Sci.* 38:531–537. doi:10.1016/j.tibs.2013.08.007.
- Cogné, M., R. Lansford, A. Bottaro, J. Zhang, J. Gorman, F. Young, H.L. Cheng, and F.W. Alt. 1994. A class switch control region at the 3' end of the immunoglobulin heavy chain locus. *Cell.* 77:737–747.
- Dudley, D.D., J. Chaudhuri, C.H. Bassing, and F.W. Alt. 2005. Mechanism and control of V(D)J recombination versus class switch recombination: similarities and differences. *Adv. Immunol.* 86:43–112. doi:10.1016/S0065-2776(04)86002-4.
- Fahey, J.L., and S. Sell. 1965. THE IMMUNOGLOBULINS OF MICE. V. THE METABOLIC (CATABOLIC) PROPERTIES OF FIVE IMMUNOGLOBULIN CLASSES. *J. Exp. Med.* 122:41–58.
- Gaudot, L., A.-S. Thomas-Claudepierre, I. Robert, E. Schiavo, V. Heyer, and B. Reina-San-Martin. 2018. Med1 deficiency alters the somatic hypermutation spectrum in murine B cells. *Eur. J. Immunol.* 48:720–723. doi:10.1002/eji.201747258.
- Ghazzoui, N., H. Issaoui, A. Saintamand, C. Oblet, C. Carrion, and Y. Denizot. 2018. The immunoglobulin heavy chain 3' regulatory region superenhancer controls mouse B1 B-cell fate and late VDJ repertoire diversity. *Blood Adv.* 2:252–262. doi:10.1182/bloodadvances.2017014423.
- Gostissa, M., B. Schwer, A. Chang, J. Dong, R.M. Meyers, G.T. Marecki, V.W. Choi, R. Chiarle, A.A. Zarrin, and F.W. Alt. 2014. *IgH* class switching exploits a general property of two DNA breaks to be joined in cis over long chromosomal distances. *Proc. Natl. Acad. Sci. U.S.A.* 111:2644–2649. doi:10.1073/pnas.1324176111.
- Kingzette, M., H. Spieker-Polet, P.C. Yam, S.K. Zhai, and K.L. Knight. 1998. Trans-chromosomal recombination within the *Ig* heavy chain switch region in B lymphocytes. *Proc. Natl. Acad. Sci. U.S.A.* 95:11840–11845.
- Laffleur, B., S.M. Bardet, A. Garot, M. Brousse, A. Baylet, and M. Cogné. 2014. Immunoglobulin genes undergo legitimate repair in human B cells not only after cis- but also frequent trans-class switch recombination. *Genes Immun.* 15:341–346. doi:10.1038/gene.2014.25.
- Le Noir, S., B. Laffleur, C. Carrion, A. Garot, S. Lecardeur, E. Pinaud, Y. Denizot, J. Skok, and M. Cogné. 2017. The *IgH* locus 3' cis-regulatory super-enhancer co-opts AID for allelic transvection. *Oncotarget.* doi:10.18632/oncotarget.14585.
- Péron, S., B. Laffleur, N. Denis-Lagache, J. Cook-Moreau, A. Tinguely, L. Delpy, Y. Denizot, E. Pinaud, and M. Cogné. 2012. AID-driven deletion causes immunoglobulin heavy chain locus suicide recombination in B cells. *Science.* 336:931–934. doi:10.1126/science.1218692.
- Pinaud, E., M. Marquet, R. Fiancette, S. Péron, C. Vincent-Fabert, Y. Denizot, and M. Cogné. 2011. The *IgH* locus 3' regulatory region: pulling the strings from behind. *Adv. Immunol.* 110:27–70. doi:10.1016/B978-0-12-387663-8.00002-8.

Reynaud, D., I.A. Demarco, K.L. Reddy, H. Schjerven, E. Bertolino, Z. Chen, S.T. Smale, S. Winandy, and H. Singh. 2008. Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat. Immunol.* 9:927–936. doi:10.1038/ni.1626.

Reynaud, S., L. Delpy, L. Fleury, H.-L. Dougier, C. Sirac, and M. Cogné. 2005. Interallelic class switch recombination contributes significantly to class switching in mouse B cells. *J. Immunol.* 174:6176–6183.

Rouaud, P., C. Vincent-Fabert, A. Saintamand, R. Fiancette, M. Marquet, I. Robert, B. Reina-San-Martin, E. Pinaud, M. Cogné, and Y. Denizot. 2013. The *IgH* 3' regulatory region controls somatic hypermutation in germinal center B cells. *J. Exp. Med.* 210:1501–1507. doi:10.1084/jem.20130072.

Saintamand, A., P. Rouaud, A. Garot, F. Saad, C. Carrion, C. Oblet, M. Cogné, E. Pinaud, and Y. Denizot. 2015a. The *IgH* 3' regulatory region governs μ chain transcription in mature B lymphocytes and the B cell fate. *Oncotarget.*

Saintamand, A., P. Rouaud, F. Saad, G. Rios, M. Cogné, and Y. Denizot. 2015b. Elucidation of *IgH* 3' region regulatory role during class switch recombination via germline deletion. *Nat Commun.* 6:7084. doi:10.1038/ncomms8084.

Stavnezer, J., J.E.J. Guikema, and C.E. Schrader. 2008. Mechanism and Regulation of Class Switch Recombination. *Annual Review of Immunology.* 26:261–292. doi:10.1146/annurev.immunol.26.021607.090248.

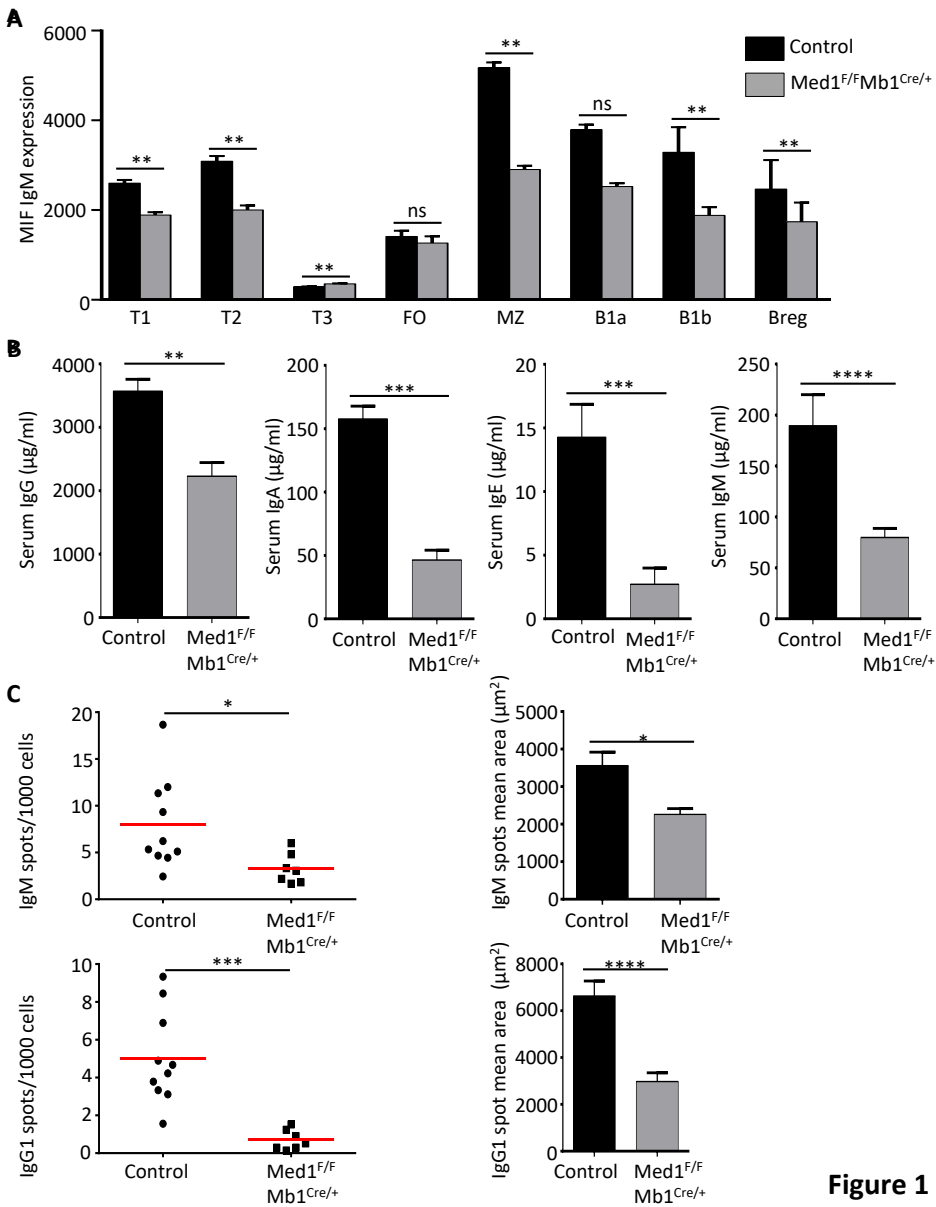
Thomas-Claudepierre, A.-S., I. Robert, P.P. Rocha, R. Raviram, E. Schiavo, V. Heyer, R. Bonneau, V.M. Luo, J.K. Reddy, T. Borggreffe, J.A. Skok, and B. Reina-San-Martin. 2016. Mediator facilitates transcriptional activation and dynamic long-range contacts at the *IgH* locus during class switch recombination. *J. Exp. Med.* 213:303–312. doi:10.1084/jem.20141967.

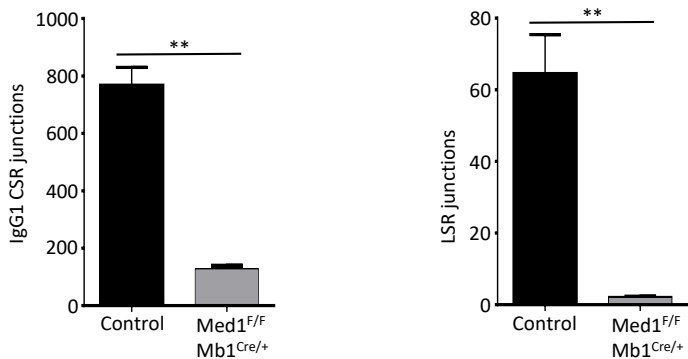
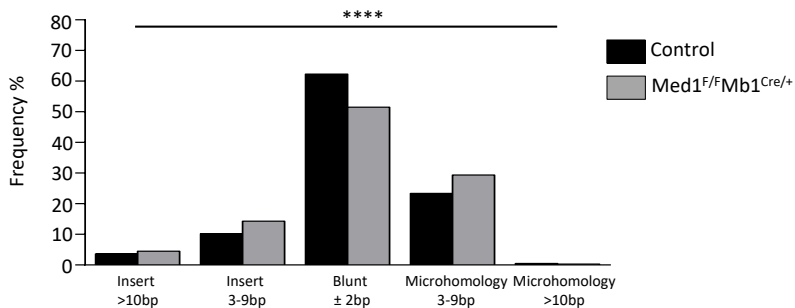
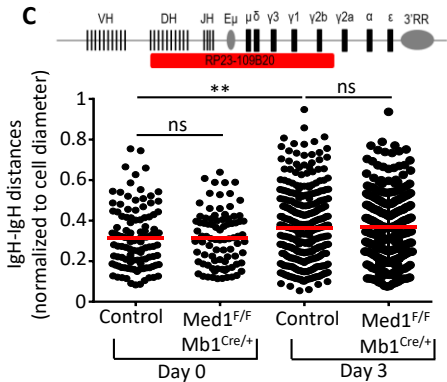
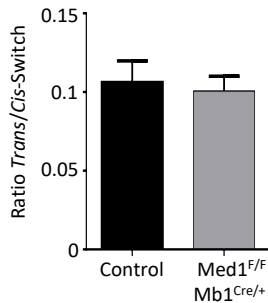
Vincent-Fabert, C., R. Fiancette, E. Pinaud, V. Truffinet, N. Cogné, M. Cogné, and Y. Denizot. 2010. Genomic deletion of the whole *IgH* 3' regulatory region (hs3a, hs1,2, hs3b, and hs4) dramatically affects class switch recombination and Ig secretion to all isotypes. *Blood.* 116:1895–1898. doi:10.1182/blood-2010-01-264689.

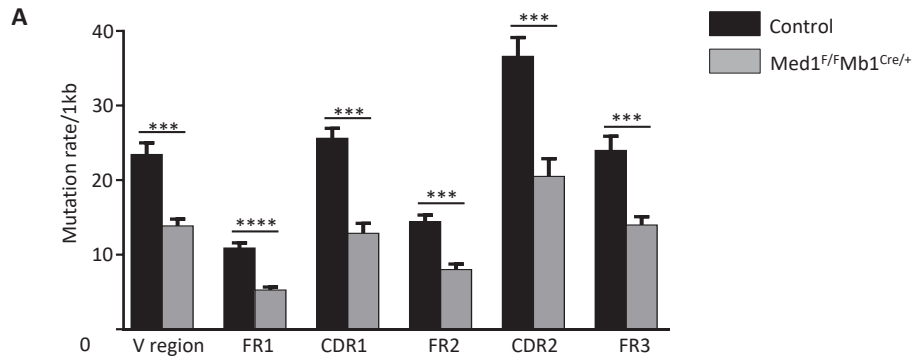
Wang, D., I. Garcia-Bassets, C. Benner, W. Li, X. Su, Y. Zhou, J. Qiu, W. Liu, M.U. Kaikkonen, K.A. Ohgi, C.K. Glass, M.G. Rosenfeld, and X.-D. Fu. 2011. Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature.* 474:390–394. doi:10.1038/nature10006.

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. doi:10.1016/j.cell.2013.03.035.

Wuerffel, R., L. Wang, F. Grigera, J. Manis, E. Selsing, T. Perlot, F.W. Alt, M. Cogné, 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.



A**B****C****D****Figure 2**



B

Control Mutation Spectra					
	A	T	C	G	Sum
A		9.4	7.2	14.8	31.4
T	3.9		5.8	4.1	13.9
C	6.6	13.0		4.5	24.1
G	19.0	4.9	6.5		30.4
					54.6

Control Mutation Spectra R					
	A	T	C	G	Sum
A		10.4	8.2	15.8	34.3
T	4.1		4.3	4.9	13.3
C	7.4	8.2		5.5	21.1
G	17.2	6.0	8.1		31.3
					52.4

Control Mutation Spectra S					
	A	T	C	G	Sum
A		6.3	4.5	11.2	22.0
T	3.9		10.0	3.1	17.0
C	5.7	25.6		2.2	33.5
G	23.4	2.2	2.0		27.5
					61.0

Med1 ^{F/F} Mb1 ^{Cre/+} Mutation Spectra					
	A	T	C	G	Sum
A		9.8	8.0	15.2	33.0
T	4.5		5.9	4.8	15.2
C	7.0	11.7		4.5	23.2
G	18.2	4.7	5.5		28.5
					51.7

Med1 ^{F/F} Mb1 ^{Cre/+} Mutation Spectra R					
	A	T	C	G	Sum
A		10.2	9.2	15.4	34.8
T	4.6		4.2	6.1	15.0
C	9.0	7.1		5.6	21.7
G	15.7	6.2	6.6		28.5
					50.2

Med1 ^{F/F} Mb1 ^{Cre/+} Mutation Spectra S					
	A	T	C	G	Sum
A		6.1	5.1	11.8	22.9
T	4.6		10.8	3.9	19.4
C	6.9	22.3		2.9	32.0
G	21.0	2.6	2.1		25.6
					57.7

Figure 3

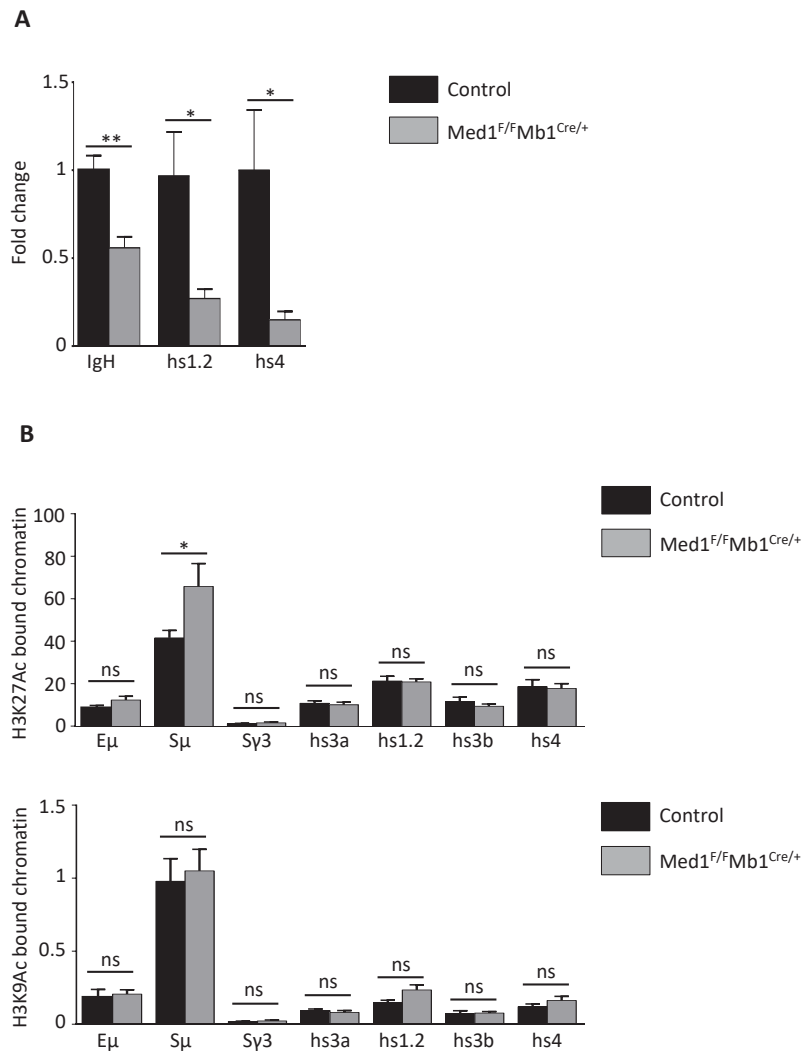
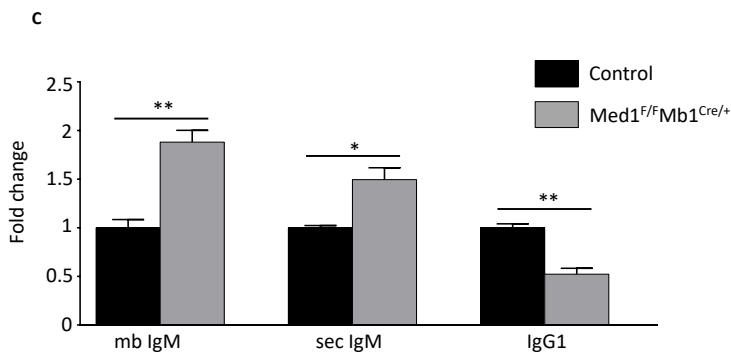
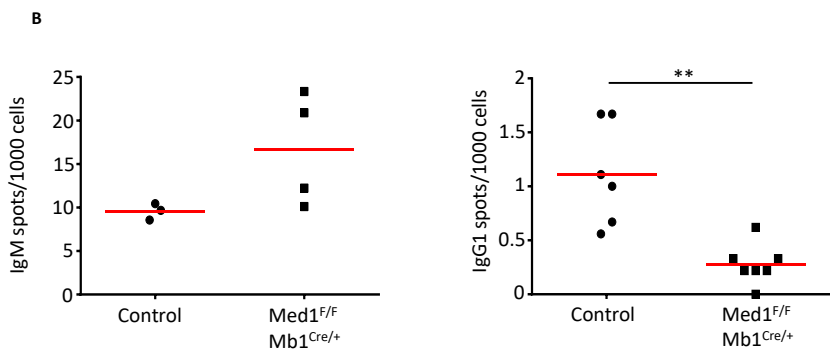
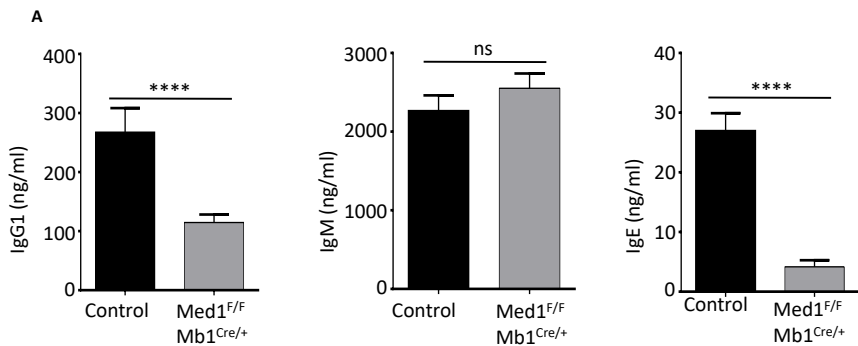


Figure 4



Article 3 : Lettre acceptée dans Cellular and Molecular Immunology

Mediator contributes to IgH locus VDJ rearrangements by promoting usage of most distal V segments

Iman Dalloul, Zeinab Dalloul, Sandrine Le Noir, François Boyer, Bernardo Reina-SanMartin, Jeanne Cook-Moreau, Michel Cogné.

Accepté en Octobre 2018

L'expression du locus d'IgH implique une configuration tridimensionnelle afin d'assurer des régulations à longue distance de diverses modifications géniques en commençant par les recombinaisons VDJ qui s'étend sur deux mégabases. Med1, une sous-unité du complexe Médiateur principalement connue pour favoriser la transcription en recrutant la polymérase II, s'est récemment avérée être impliquée dans des événements de recombinaison affectant la commutation de classe du locus IgH dans les cellules B matures.

En vérifiant le répertoire VDJ du locus IgH en utilisant la technique du séquençage à haut débit, nous montrons maintenant que Med1 a également un rôle vraisemblable dans les réarrangements précoces qui permettent d'établir le répertoire des cellules B. Med1 favorise ainsi l'expression des gènes distaux V, situés à plus de 500 kb de l'enhancer intronique E μ .

Nos résultats indiquent donc que Med1 contribue ainsi globalement aux réarrangements géniques à longue distance dans le locus IgH, qu'ils soient initiés par AID dans les lymphocytes B matures ou par les RAG dans les progéniteurs des cellules B.



CORRESPONDENCE

Mediator contributes to IgH locus VDJ rearrangements by promoting usage of most distal V segments

Iman Dalloul¹, Zeinab Dalloul¹, Sandrine Le Noir¹, François Boyer¹, Bernardo Reina-San-Martin², Jeanne Cook-Moreau¹ and Michel Cogné¹

Keywords: B cells; mediator; VDJ repertoire

Cellular & Molecular Immunology _____; <https://doi.org/10.1038/s41423-018-0175-z>

INTRODUCTION

The immunoglobulin heavy (*IgH*) chain locus includes a large cluster of V, D and J segments spanning thousands of kilobases in mammals and its expression implies long-range 3D-regulation of various gene modifications.¹ V(D)J recombination occurs during early B-cell ontogeny.² It is regulated by multiple transcription factors and *cis*-regulatory elements, and most notably by the intronic 5' E μ element located in the J_H to C μ intron.³ The intergenic control region 1 (IGCR1) located between the V_H and D gene clusters inhibits rearrangement of the most D_H-proximal V_H gene segments, while promoting usage of distal V_H gene segments.⁴ Other downstream elements of the locus, flanking its 3' end might have more subtle roles: while the large 3' regulatory region (3'RR) shows strong internal synergies² and behaves as a superenhancer in mature B cells. Its deletion showed no influence on the VDJ repertoire and rather increased transcription of unrearranged V segments.^{3–5} Insulators binding CTCF farther downstream of the 3'RR were shown, however, to promote usage of distal V_H segments.⁶ Altogether, there are strong indications that elements involved in long-distance regulation of transcription might also influence VDJ recombination. In-frame splicing of VDJ exons is also needed for full Ig expression.⁷ Given the functional importance of gene architecture and long-range interactions for the *IgH* locus physiology, we explored the dependence of the VDJ repertoire on Med1, a known actor of long-range enhancer interactions. By checking the repertoire using high-throughput sequencing, we now show that Med1 also has a role in early rearrangements and promotes expression of distal V_H genes. Med1 deficiency resulted in a significant repertoire bias, with decreased usage of distal V segments, together with decreased usage of the upstream JH1 and JH2 segments.

Procedures were approved by the local ethics committee review board and conducted according to the European guidelines for animal experimentation. To specifically deplete Med1 in the B-cell lineage, Med1-floxed mice (Med1^{F/F}) were bred with Mb1Cre/ + knock-in mice, a conditional ablation known to globally preserve B-cell development.⁸ Single cell suspensions of splenocytes from Med1^{F/F} mb1-cre and control *WT* C57Bl6/SV129 mice were stained with CD19-APCH7, CD5-FITC, CD21-PE, CD23-PECy7, IgD-BV421 and IgM-APC antibodies for flow cytometry. As previously described, normally abundant peripheral B cells were present

upon deletion of Med1, with an increased ratio of marginal zone vs follicular splenic B cells (Fig. 1a). Detailed analysis of V regions from sequenced *IgH* transcripts revealed a significant alteration of V and J_H usage (Fig. 1b). The VDJ repertoire was specifically studied in RNA extracted from mature B cells after dissociation of Peyer's patches from Med1-deficient mice. RACE-PCR amplicons were sequenced on an Illumina Miseq sequencer. Data were analyzed with the international ImMunoGeneTics IMGT/HighV-QUEST Web portal, and IMGT-ONTOLOGY concepts. V_H segments showed increased usage of proximal genes (i.e. those located less than 500 bp upstream of J_H). J_H distribution was also altered, with increased use of those J_H closest to the E μ enhancer (J_H4 and J_H3) and decreased usage of the more distant J_H2 and J_H1.

Mediator is a four-module complex with several important functions in binding transcription factors to RNA polymerase II (Pol II). It regulates Pol II C-terminal domain phosphorylation and Pol II status regarding transcription initiation, pausing or elongation.⁹ While not mandatory for transcription and lacking in some Mediator complexes, the Med1 subunit is characterized by specific and direct interaction with several transcriptional activators binding enhancer elements. It is involved in functional and physical loops between promoter and enhancer elements that stimulate transcription in activated cells.¹⁰ A single report described Med1 influence on gene recombination in activated mature B cells, where it promoted the long-distance interactions of IgH enhancers with target switch-regions and globally enhanced CSR to all Ig classes by 2 to 3-fold.⁸ VDJ recombination also involves long-distance interactions promoted by enhancer elements which control locus contraction.^{11,12}

While VDJ rearrangement is contemporary to early B-cell development and occurs in the absence of B-cell activation, we explored whether this process was under the influence of Med1 as well. Although Med1 deficiency did not affect peripheral B-cell populations as previously described, checking the Ig VDJ repertoire from Med1-deficient B cells revealed a decreased usage of the J_H segments that are more distal from the E μ enhancer (J_H1 and J_H2), together with decreased usage of the distal cluster of V_H segments. The variable exon bias was reminiscent of anomalies previously reported in several models deficient for factors that contribute to early accessibility and chromosomal looping during the so-called 'IgH locus contraction'.^{12–15} Med1 thus most likely

¹UMR CNRS 7276, INSERM U1262, Université de Limoges, Limoges 87025, France and ²INSERM U1258, UMR CNRS 7104, Institut de Genetique et de Biologie Moleculaire et Cellulaire, Université de Strasbourg, Illkirch 67400, France
Correspondence: Michel Cogné (cogne@unilim.fr)

Received: 4 October 2018 Accepted: 8 October 2018

Published online: 07 November 2018

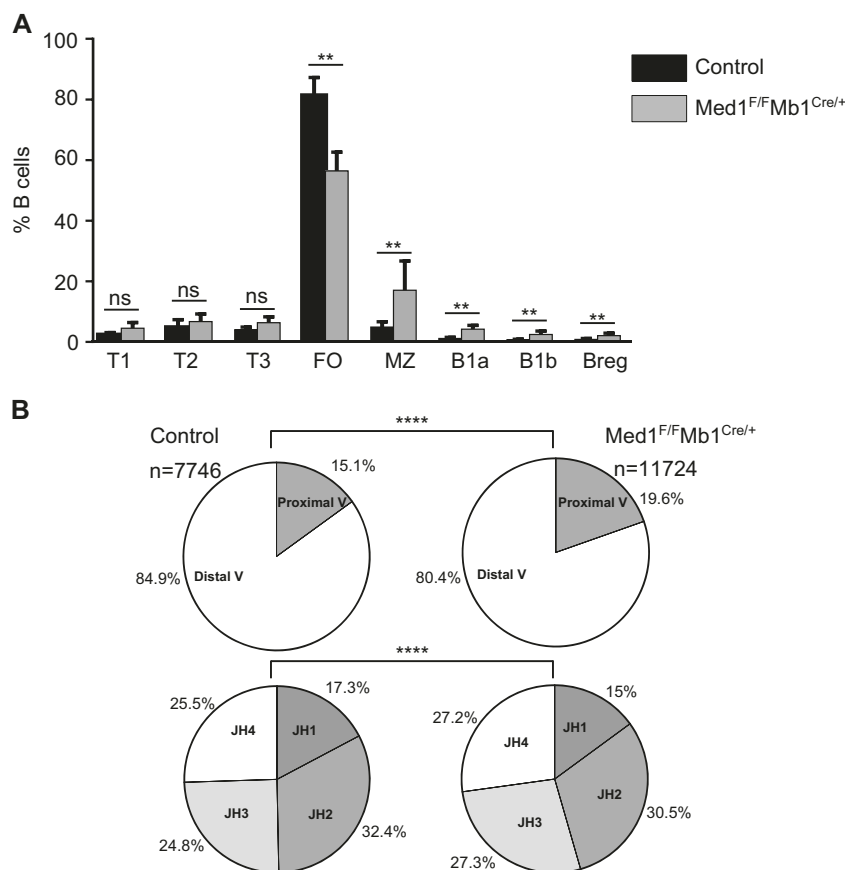


Fig. 1 B-cell populations and VDJ recombination in Med1^{F/F}Mb1^{Cre/+} mice. **a** Spleen B-cell populations were studied by flow cytometry on wt and Med1^{F/F}mb1^{Cre/+} mice. Data represent mean ± SEM from 6 wt and 6 Med1^{F/F}Mb1^{Cre/+} mice. Mann–Whitney test was used for significance. **b** VDJ repertoire was studied in Peyer’s patches from Med1^{F/F}Mb1^{Cre/+} mice after NGS sequencing using a C α specific reverse primer. Proximal and distal V_H (top) and J_H H (bottom) were studied on 7746 clonotypes in control mice (n = 7) and 11724 clonotypes for Med1^{F/F}Mb1^{Cre/+} mice (n = 11)

contributes to long-distance gene rearrangements in the IgH locus, whether initiated by AID in mature B cells or by RAG in B-cell progenitors assembling their primary VDJ repertoire.

ACKNOWLEDGEMENTS

This work was supported by grants from Association pour la Recherche sur le Cancer (PGA120150202338), ANR grant Ig-MemImpact 16-CE15-0019-01) and ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the frame program Investissements d’Avenir ANR-10-IDEX-0002-02. We thank Mehdi Alizadeh for his help with MiSeq sequencing.

AUTHOR CONTRIBUTIONS

M.C., J.C.-M. designed the study. I.D., S.L.N. and Z.D. performed experiments. F.B. analysed Ig repertoire data. B.R.-S.-M. provided mice and participated in manuscript writing.

ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

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

REFERENCES

- Pinaud, E. et al. The IgH locus 3’ regulatory region: pulling the strings from behind. *Ad. Immunol.* **110**, 27–70 (2011).

- Chauveau, C., Pinaud, E. & Cogne, M. Synergies between regulatory elements of the immunoglobulin heavy chain locus and its palindromic 3’ locus control region. *Eur. J. Immunol.* **28**, 3048–3056 (1998).
- Morvan, C. L., Pinaud, E., Decourt, C., Cuvillier, A. & Cogné, M. The immunoglobulin heavy-chain locus hs3b and hs4 3’ enhancers are dispensable for VDJ assembly and somatic hypermutation. *Blood* **102**, 1421–1427 (2003).
- Rouaud, P. et al. Enhancers located in heavy chain regulatory region (hs3a, hs1,2, hs3b, and hs4) are dispensable for diversity of VDJ recombination. *J. Biol. Chem.* **287**, 8356–8360 (2012).
- Braikia, F. Z. et al. Developmental switch in the transcriptional activity of a long-range regulatory element. *Mol. Cell. Biol.* **3**, 3370–3380 (2015).
- Guo, C. et al. CTCF-binding elements mediate control of V(D)J recombination. *Nature* **477**, 424–430 (2011).
- Ashi M. O., et al. Physiological and druggable skipping of immunoglobulin variable exons in plasma cells. *Cell Mol Immunol* 2018. <https://doi.org/10.1038/s41423-018-0160-6>
- Thomas-Claudepierre, A.-S. et al. Mediator facilitates transcriptional activation and dynamic long-range contacts at the IgH locus during class switch recombination. *J. Exp. Med.* **213**, 303–312 (2016).
- Allen, B. L. & Taatjes, D. J. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 155–166 (2015).
- Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
- Fuxa, M. et al. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev.* **18**, 411–422 (2004).
- Verma-Gaur, J. et al. Noncoding transcription within the Igh distal VH region at PAIR elements affects the 3D structure of the Igh locus in pro-B cells. *Proc. Natl Acad. Sci. USA* **109**, 17004–17009 (2012).
- Xu, C.-R., Schaffer, L., Head, S. R. & Feeney, A. J. Reciprocal patterns of methylation of H3K36 and H3K27 on proximal vs. distal IgVH genes are modulated by IL-7 and Pax5. *Proc. Natl Acad. Sci. USA* **105**, 8685–8690 (2008).

14. Nutt, S. L., Heavey, B., Rolink, A. G. & Busslinger, M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401**, 556–562 (1999).

15. Su, I.-H. et al. Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat. Immunol.* **4**, 124–131 (2003).

Discussion et perspectives

Le locus de la chaîne lourde des immunoglobulines (IgH) regroupe des gènes qui sont étroitement régulés tout au long du développement et la différenciation des cellules B. Quand une cellule B rencontre un antigène, le locus IgH peut subir plusieurs processus de diversification qui affectent à la fois les gènes variables (V) et constants (C). Ces processus dépendent d'un ensemble d'éléments trans ou cis-activateurs, mais leur « chef d'orchestre » au sein du locus semble vraiment être constitué par la région régulatrice 3'RR. En tant que régulateur de l'expression du locus IgH, ce *super-enhancer* contrôle plusieurs événements à divers stades de la maturation des cellules B naïves aux cellules mémoires, jusqu'au stade des cellules plasmocytaires. Le locus IgH humain portant une duplication ancestrale des gènes constants, il porte aussi deux exemplaires de « 3'RR », en aval de chacun des deux gènes C α . Les deux régions régulatrices 3'RR humaines montrent une forte similarité d'architecture avec la région 3'RR murine avec une structure palindromique flanquant l'*enhancer* hs1-2. Les deux régions 3'RR humaines présentent des régions Like-switch (LS) (5 régions LS dans la 3'RR1 et 6 LS dans la 3'RR2) qui ressemblent aux régions switch du fait de leur structure répétée, mais où nous n'avons pas observé une aussi grande densité de motifs cibles d'AID (WRCY ou RGYW, sites préférentiels de déamination par cet enzyme). Ces régions LS hautement répétitives ont chez l'Homme une taille entre 0.5 à 2 kb, ce qui est proche de la longueur de S α 2 et elles sont donc plus organisées et longues que les 11 régions LS préalablement décrites dans la 3'RR de la souris (taille entre 0.5 à 1 kb). Notre travail a montré par un séquençage à haut débit des jonctions de recombinaison, que même si les motifs de ciblage AID sont plus rares, ils semblent se situer à proximité des points de cassure, ce qui signe donc l'implication significative d'AID dans le processus des cassures LSR survenant au niveau du *super-enhancer* 3'RR. De même, la transcription des régions ciblées et leur richesse en G sont des éléments de nature à y favoriser la formation de structures R-loops, extrêmement susceptibles aux lésions de l'ADN. AID pourrait donc attaquer, comme pour la CSR, au niveau de ces régions LS et des R-loops associées, initiant des cassures simple brin, puis double brin, et aboutissant à la délétion de l'ensemble des gènes constants du locus IgH. Comme chez la souris (Péron et al., 2012a), ce processus conduirait à la mort des cellules B humaines. Notre étude montre aussi chez l'Homme que ce type de recombinaison entre la région S μ et les 3'RR1 et 3'RR2 (jonction S μ -3'RR) existe dans les cellules B activées in-vitro. Ce phénomène est, comme pour la CSR, AID-dépendant marqué par la quasi absence de ce type de jonction dans les cellules B AID déficientes (Muramatsu et al., 2000). Ces résultats sont corrélés avec ceux de la souris qui

montrent que les cellules B activées subissent la LSR seulement dans des cellules B AID-proficientes (Péron et al., 2012a). Les cellules B mémoires activées mais aussi les plasmablastes triés à partir des amygdales montrent un fort signal LSR. Ces jonctions ne sont au contraire presque pas détectées ni dans le sang, ni dans les cellules B au repos ni dans les plasmocytes au niveau de la moelle osseuse. Klein et ses collègues ont montré que le sang humain contient 40% de cellules B mémoires (Klein et al., 1998), et comme la LSR induit la mort des cellules B, cela suggère que ce type de recombinaison n'apparaisse que transitoirement dans les cellules B activées et que AID attaque toujours l'allèle fonctionnel (ou bien agit pour la LSR comme le plus souvent pour la CSR, de façon bi-allélique). Les cassures générées pendant la LSR devraient, comme pour la CSR, potentiellement être réparées par une des deux voies de réparation NHEJ ou A-EJ. Durant la CSR l'étape de résolution est majoritairement médiée par la voie de réparation classique NHEJ qui génère des jonctions CSR directe et de courte microhomologie ; cependant, d'autres études suggèrent que la résolution peut également être médiée par la voie de réparation alternative A-EJ qui génèrent des jonctions avec des longues microhomologies et des insertions (Pan-Hammarström et al., 2005; Soulas-Sprauel et al., 2007; Yan et al., 2007). L'analyse de la structure des jonctions LSR humains montre par comparaison à la CSR une augmentation de l'utilisation des microhomologies et des insertions corrélées avec le mécanisme de réparation utilisée par la voie A-EJ.

Dans nos études, avec le handicap que pouvait représenter la lourdeur des tests de séquençage haut-débit nécessaires pour avoir une appréciation semi-quantitative du phénomène de LSR, nous n'avons pas pu objectiver de variation statistiquement significative du rapport LSR/CSR en faisant varier les conditions de stimulations (notamment par des variations des cytokines en présence). Actuellement, il n'existe donc pas aujourd'hui d'éléments permettant de comprendre s'il existe des « checkpoints » et des conditions précises de régulation du destin de la cellule B qui l'orienteront plus spécifiquement vers CSR et survie ou vers LSR et apoptose.

C'est dans ce cadre que nous avons souhaité tester le rôle de facteurs potentiellement impliqués dans la formation des interactions géniques à longue distance au sein des locus, tel que le complexe Médiateur. Comme la sous-unité MED1 du complexe Médiateur a montré qu'elle stimule la CSR et qu'elle influence aussi qualitativement la SHM, deux mécanismes étroitement contrôlés par la région 3'RR (Gaudot et al., 2018; Thomas-Claudepierre et al., 2016), nous nous sommes particulièrement intéressés à étudier l'impact de sa délétion sur le locus IgH. Notre étude a révélé que la SHM est en fait elle aussi quantitativement dépendante de MED1, de même que l'expression du BCR et la sécrétion des Ig par les plasmocytes. Enfin, les jonctions

LSR chez la souris déficiente en la sous-unité MED1 du complexe médiateur sont fortement diminuées. Ceci suggère un rôle crucial de cette sous-unité dans tous les mécanismes dépendants du *super-enhancer* 3'RR, et en particulier dans le mécanisme de la LSR.

Dans les cellules B murines au repos la sous-unité MED1 est recrutée au niveau des deux activateurs transcriptionnels E μ et la 3'RR favorisant la formation d'une boucle entre ces deux régions. Après activation de ces cellules B, le complexe médiateur peut également se fixer au niveau de la région switch activatrice permettant son interaction avec E μ et la 3'RR, ce qui induit le recrutement de l'ARN polymérase II et l'initiation de la transcription germinale au niveau de la région switch (Thomas-Claudepierre et al., 2016). Notre laboratoire a décrit que la région 3'RR (*hs1-2* et *hs4*) murine est transcrite, ce qui est un prérequis pour la LSR, comme pour la CSR, pour permettre l'attaque d'AID au niveau de ces régions cibles (Péron et al., 2012a). La diminution drastique des jonctions LSR dans les souris déficientes en la sous-unité MED1 pourrait s'expliquer par la délétion de la sous-unité MED1 dans les cellules B avec pour conséquence perturbation de la formation de la boucle entre E μ et la 3'RR (Thomas-Claudepierre et al., 2016) et inhibition de la transcription germinale. Nos résultats vont dans ce sens car nous avons observé une diminution de la transcription de la région 3'RR chez ces souris.

Outre son effet sur la diminution de la LSR, nous avons observé que cette délétion est accompagnée par une diminution de la sécrétion de toutes les classes d'Ig *in-vivo* comme *in-vitro*. De façon intéressante, on note aussi que ce knock-out affecte l'expression du locus IgH (et aussi la transcription en eRNA du *super-enhancer* 3'RR) sans avoir pour autant d'impact sur les marques épigénétiques du locus IgH, ce qui renforce l'idée que son action est purement « architecturale » sur la conformation du locus et le recrutement de Pol II. En accord avec les résultats de Thomas-Claudepierre et al., nous avons trouvé un discret biais de développement des cellules B au niveau de la rate chez les souris MED1 KO qui se traduit par une diminution des cellules B folliculaires et une augmentation du pourcentage des cellules B marginales. En plus du défaut de CSR, nous avons observé des caractéristiques supplémentaires, notamment une diminution de la densité de l'expression du BCR à la surface de toutes les populations des cellules B spléniques (Thomas-Claudepierre et al., 2016).

La sous-unité MED1 est impliquée dans les interactions à longues distances du locus IgH et dans le mécanisme de CSR. Contrairement à la 3'RR qui favorise les interactions inter-alléliques du locus IgH, la sous-unité MED1 ne semble pas être impliquée. Comme notre

laboratoire s'est récemment intéressé aux recombinaisons inter-chromosomiques au cours de la CSR et a montré que la 3'RR était nécessaire au rapprochement des deux allèles du locus IgH qui précède le mécanisme de *trans*-switch (Le Noir et al., 2017b), nous avons cherché à savoir si MED1 pouvait être encore plus crucial pour ces interactions très distantes entre chromosomes différents, qu'il ne l'est aux interactions intra-chromosomiques. Nos résultats montrent en fait qu'il n'en est rien et qu'en l'absence de MED1, le classe switching est affecté de façon équivalente, tant pour le cis-CSR que le trans-CSR, que les interactions à longue distance soient donc en intra ou en inter-allélique.

Outre l'implication de la 3'RR dans les interactions à longues distances avec E μ mais aussi avec la région S μ et la région switch acceptrice, il a été largement étudié le rôle de certaines protéines qui peuvent intervenir dans la formation d'une boucle entre la région 3'RR et les activateurs *hs* situés dans la région VDJ. Volpi et ses collègues ont montré que la délétion des isolateurs *hs5-hs7* qui servent de « borne 3' » à la 3'RR du locus IgH et qui contiennent des sites de fixation des CTCF favorisent le réarrangement des VH proximaux par rapport aux VH distaux dans les cellules pro-B (Volpi et al., 2012). Il est à noter que ces sites pour CTCF se trouvent également au voisinage des gènes VH, et ces auteurs ont postulé la formation de boucles associant les sites *hs5-hs7* et ceux des régions VH en prélude aux réarrangements VDJ. L'observation d'une altération similaire chez les souris délétées en la sous-unité MED1, avec augmentation des réarrangements des VH proximaux et diminution des réarrangements des VH distaux peut donc à nouveau être interprété en termes de boucles entre des régions distantes du locus, et notamment sa région 3' régulatrice.

La diminution de la SHM et la diminution de transcription du locus IgH est pour sa part aussi une « phénocopie » de la baisse de SHM observée dans plusieurs modèles knock-out portant des altérations de la 3'RR (Le Noir et al., 2017a; Rouaud et al., 2013; Saintamand et al., 2016).

Perspectives

La description du phénomène de la LSR tout d'abord chez la souris puis plus récemment chez l'Homme pourrait constituer une nouvelle piste pour mieux comprendre l'homéostasie des populations lymphocytaires B, en physiologie comme en pathologie. Nous avons déjà tenté en vain de rechercher un éventuel phénomène d'augmentation de la LSR dans certaines maladies caractérisées par des lymphopénies B sévères telles que le sepsis. Nos premières analyses des jonctions LSR faites sur le sang des patients septiques ne montrent pas d'augmentation statistiquement significative de la LSR chez les patients hospitalisés et traités. Cependant, il est

clair que dans ces situations, les tests sont faits après la période critique, et ne se prêtent donc pas à l'observation chez l'Homme d'un phénomène transitoire, qui pourrait ne se solder que par une diminution du nombre des cellules B dans le sang. Pour des raisons tant pratiques qu'éthiques, il sera difficile d'obtenir des échantillons humains correspondant à la phase aigüe et la plus dramatique du sepsis. Nous avons donc initié en parallèle une étude du sepsis chez la souris, un utilisant le modèle classique de perforation/ligation du cæcum. De façon préliminaire, des signaux LSR forts sont observés dans ces conditions, mais des signaux CSR élevés coexistent et des mises au point sont encore demandées afin de pouvoir tirer des conclusions fiables.

Une des hypothèses faites à propos du rôle de la LSR dans l'homéostasie B est aussi son rôle éventuel dans l'élimination de cellules auto-réactives. Pour cette raison, il sera aussi intéressant d'étudier le phénomène sur des échantillons d'animaux présentant des manifestations d'auto-réactivité ou chez des personnes ayant des maladies auto-immunes. L'étude de la LSR sur ces échantillons pourraient permettre de mieux comprendre ce phénomène.

A l'inverse, un travail en cours dans le laboratoire concerne la plus grande abondance des IgM ayant une auto-réactivité chez des animaux présentant un défaut de LSR induit par un knock-out (de la région $S\mu$ ou de la région 3'RR).

D'autre part, l'analyse de la structure des jonctions LSR montrent une éventuelle implication majeure de la voie alternative A-EJ dans la réparation de ces jonctions. Donc il sera aussi intéressant d'étudier ce phénomène dans des échantillons mais aussi dans des souris déficientes en des protéines impliquées dans ces deux voies de réparation des cassures.

Enfin nos données sur le KO MED1 sont alléchantes, car il s'agit du premier exemple de situation dans laquelle apparaît une atteinte différentielle de la LSR, supérieure à celle observée pour la CSR. La LSR serait donc tout particulièrement Médiateur-dépendante. Concernant l'implication de la sous-unité MED1 du complexe Médiateur dans les autres remaniements qui touchent le locus IgH pendant le développement des cellules B marqué spécifiquement par un défaut de la CSR, mais aussi un changement dans le profil de mutation qui touche la région variable, nous envisageons de regarder si le défaut de sécrétion des immunoglobulines est dû à un défaut au niveau des cellules plasmocytaires. Comme nos premiers résultats ont montré un défaut de réarrangement VDJ dans les cellules B switchés en IgA, donc il sera aussi important de regarder l'effet de la délétion de la sous-unité MED1 sur la recombinaison VDJ dans les cellules pré-B triées de la moelle, là où les recombinaisons VDJ débutent mais aussi de regarder

les interactions à longues distances qui pourraient être perturbées au sein du locus IgH et qui peuvent être la cause des phénotypes observés dans ces souris.

Annexe 1: (Article)

CSReport: A New Computational Tool Designed for Automatic Analysis of Class Switch Recombination Junctions Sequenced by High-Throughput Sequencing.

François Boyer, Hend Boutouil, **Iman Dalloul**, Zeinab Dalloul, Jeanne Cook-Moreau, Jean-Claude Aldigier, Claire Carrion, Bastien Herve, Erwan Scaon, Michel Cogné and Sophie Péron.

Article accepté à The journal of Immunology en mai 2017.

Dans cette publication nous avons développé un outil bio-informatique que nous avons nommé CSReport et qui permet d'analyser les jonctions CSR chez la souris mais aussi chez l'Homme en alignant les reads obtenus après le séquençage Ion proton NGS avec le locus IgH utilisé comme une séquence de référence dans ce programme. CSReport identifie les points de cassure qui se produisent durant la CSR au niveau de S μ et d'une région switch acceptrice mais elle nous permet aussi d'analyser la structure de ces jonctions (microhomologies et insertions). Nous avons ultérieurement utilisé cet outil pour analyser les jonctions LSR chez l'Homme et chez la souris.

Ma contribution dans cet article a consisté en la réalisation des stimulations in-vitro sur les cellules mononuclées du sang humains PBMC. J'ai fait les PCR qui étudient le switch vers IgA que j'ai utilisé pour préparer les bibliothèques pour le séquençage Ion proton NGS.



Selection Antibiotics

Keep what you want, lose the rest!

Blasticidin, Zeocin™, Puromycin, G418, Hygromycin B Gold and Phleomycin

InvivoGen



This information is current as of October 14, 2018.

CSReport: A New Computational Tool Designed for Automatic Analysis of Class Switch Recombination Junctions Sequenced by High-Throughput Sequencing

François Boyer, Hend Boutouil, Iman Dalloul, Zeinab Dalloul, Jeanne Cook-Moreau, Jean-Claude Aldigier, Claire Carrion, Bastien Herve, Erwan Scaon, Michel Cogné and Sophie Péron

J Immunol 2017; 198:4148-4155; Prepublished online 17 April 2017;
doi: 10.4049/jimmunol.1601924
<http://www.jimmunol.org/content/198/10/4148>

Supplementary Material <http://www.jimmunol.org/content/suppl/2017/04/16/jimmunol.1601924.DCSupplemental>

References This article **cites 16 articles**, 6 of which you can access for free at:
<http://www.jimmunol.org/content/198/10/4148.full#ref-list-1>

Why *The JI*? Submit online.

- **Rapid Reviews! 30 days*** from submission to initial decision
- **No Triage!** Every submission reviewed by practicing scientists
- **Fast Publication!** 4 weeks from acceptance to publication

**average*

Subscription Information about subscribing to *The Journal of Immunology* is online at:
<http://jimmunol.org/subscription>

Permissions Submit copyright permission requests at:
<http://www.aai.org/About/Publications/JI/copyright.html>

Email Alerts Receive free email-alerts when new articles cite this article. Sign up at:
<http://jimmunol.org/alerts>

The Journal of Immunology is published twice each month by
The American Association of Immunologists, Inc.,
1451 Rockville Pike, Suite 650, Rockville, MD 20852
Copyright © 2017 by The American Association of
Immunologists, Inc. All rights reserved.
Print ISSN: 0022-1767 Online ISSN: 1550-6606.



CSReport: A New Computational Tool Designed for Automatic Analysis of Class Switch Recombination Junctions Sequenced by High-Throughput Sequencing

François Boyer,^{*,†,1} Hend Boutouil,^{*,†,1} Iman Dalloul,^{*,†} Zeinab Dalloul,^{*,†}
Jeanne Cook-Moreau,^{*,†} Jean-Claude Aldigier,^{*,†} Claire Carrion,^{*,†} Bastien Herve,[‡]
Erwan Scaon,[‡] Michel Cogné,^{*,†,§} and Sophie Péron^{*,†}

B cells ensure humoral immune responses due to the production of Ag-specific memory B cells and Ab-secreting plasma cells. In secondary lymphoid organs, Ag-driven B cell activation induces terminal maturation and Ig isotype class switch (class switch recombination [CSR]). CSR creates a virtually unique *IgH* locus in every B cell clone by intrachromosomal recombination between two switch (S) regions upstream of each C region gene. Amount and structural features of CSR junctions reveal valuable information about the CSR mechanism, and analysis of CSR junctions is useful in basic and clinical research studies of B cell functions. To provide an automated tool able to analyze large data sets of CSR junction sequences produced by high-throughput sequencing (HTS), we designed CSReport, a software program dedicated to support analysis of CSR recombination junctions sequenced with a HTS-based protocol (Ion Torrent technology). CSReport was assessed using simulated data sets of CSR junctions and then used for analysis of $S\mu$ - $S\alpha$ and $S\mu$ - $S\gamma 1$ junctions from CH12F3 cells and primary murine B cells, respectively. CSReport identifies junction segment breakpoints on reference sequences and junction structure (blunt-ended junctions or junctions with insertions or microhomology). Besides the ability to analyze unprecedentedly large libraries of junction sequences, CSReport will provide a unified framework for CSR junction studies. Our results show that CSReport is an accurate tool for analysis of sequences from our HTS-based protocol for CSR junctions, thereby facilitating and accelerating their study. *The Journal of Immunology*, 2017, 198: 4148–4155.

B cells ensure humoral immune responses due to their capacity to recognize a wide variety of Ags via their BCR for Ag. BCR is the association of an Ig, composed of two

identical IgH chains covalently bound to two identical IgL chains, which are associated with an $Ig\alpha/Ig\beta$ heterodimer at the B cell membrane. In mature B cells, in secondary lymphoid organs, Ag binding by the BCR with T cell–dependent coactivation signaling induces Ag-driven activation of B cells, which then undergo terminal maturation, ensuring production of Ag-specific memory B cells and Ab-secreting plasma cells. Two molecular events absolutely necessary for efficient humoral responses against Ag occur, as follows: Ig isotype class switch (class switch recombination [CSR]) and somatic hypermutation (SHM) (1, 2). CSR is an intrachromosomal recombination that occurs in the *IgH* locus between two switch (S) regions located upstream of each C segment encoding the C region of the Ig. SHM consists in the introduction of mutations in the V regions of *IgH* ($V_{H}D_{H}J_{H}$) and *IgL* ($V_{L}J_{L}$) loci, which recognize the Ag. CSR enables synthesis of different Ig isotypes, each characterized by distinct effector functions, without affecting the antigenic specificity of Ig. SHM permits the selection of B cell clones with high affinity for Ag. CSR is divided into three steps, as follows: 1) germline transcription of targeted S regions; 2) DNA lesions induced by the enzyme activation-induced cytidine deaminase (1, 3) leading to double-strand breaks (DSB) in DNA; and finally 3) DNA repair mainly by the nonhomologous end joining (NHEJ) pathway.

Analysis of CSR junctions has been used in basic and clinical research (for review, see Ref. 4) because it provides insight into the amount of class-switched B cells in samples and documents the predominant CSR molecular pathway. In response to Ag, B cells expand and CSR creates a virtually unique *IgH* locus in every B cell clone, and CSR junctions can be amplified by PCR in a given sample. This is especially needed to evaluate the potential alterations of CSR in various immune defects. Structure of the repair

*Université de Limoges, Contrôle de la Réponse Immune B et Lymphoproliférations, UMR 7276, F-87000 Limoges, France; †CNRS, Contrôle de la Réponse Immune B et Lymphoproliférations, UMR 7276, F-87000 Limoges, France; ‡Université de Limoges, plateforme Biologie Intégrative Santé Chimie Environnement, F-87000 Limoges, France; and §Institut Universitaire de France, F-75000 Paris, France

¹F.B. and H.B. contributed equally to this work.

ORCID: 0000-0002-0351-685X (F.B.); 0000-0003-3590-1602 (J.C.-M.); 0000-0003-3221-6319 (C.C.); 0000-0002-8519-4427 (M.C.); 0000-0002-0056-3271 (S.P.).

Received for publication November 10, 2016. Accepted for publication February 19, 2017.

This work was supported by grants from Conseil Régional du Limousin, Association pour la Recherche sur le Cancer, and Ligue contre le Cancer (Comité de la Haute-Vienne, Comité de la Corrèze). F.B. is supported by Fondation Partenariale de l'Université de Limoges and Association Limousine pour l'Utilisation du Rein Artificiel à Domicile. H.B. is supported by Région Limousin and Université de Limoges. M.C. is supported by Institut Universitaire de France and Fondation pour la Recherche Médicale.

The sequences presented in this article have been submitted to the European Molecular Biology Laboratory–European Bioinformatics Institute European Nucleotide Archive (<http://www.ebi.ac.uk/ena>) under the accession number ERP019632.

See related articles in this issue: Langerak et al. (*J. Immunol.* 198, 3765; DOI: <https://doi.org/10.4049/jimmunol.1602050>) and IJspeert et al. (*J. Immunol.* 198, 4156; DOI: <https://doi.org/10.4049/jimmunol.1601921>).

Address correspondence and reprint requests to Dr. Sophie Péron, Université de Limoges, CNRS UMR 7276, Centre de Biologie et de Recherche en Santé, Faculté de Médecine, 2 rue du Docteur Marcland, Limoges 87000, France. E-mail address: sophie.peron@inserm.fr

The online version of this article contains supplemental material.

Abbreviations used in this article: CSR, class switch recombination; DSB, double strand break; HTS, high-throughput sequencing; NHEJ, nonhomologous end joining; SHM, somatic hypermutation; S region, switch region.

Copyright © 2017 by The American Association of Immunologists, Inc. 0022-1767/17/\$30.00

junction between donor and acceptor segments indeed reflects the DNA DSB repair pathway used during CSR: it has been shown that a recombination junction between two S regions may be repaired by alternative end joining DNA (5, 6) in the case of defective NHEJ. CSR recombination junction analysis protocols (6, 7) until now relied on cloning and subsequent sequencing by Sanger's method of individual products obtained by a single PCR or a nested PCR designed to amplify DNA fragments from switch junctions. Alignment of sequences with germline reference sequences assigns breakpoints on the *IgH* locus for each donor and acceptor segments involved, characterizes individual junction structure, and identifies mutations around the CSR recombination sites.

To overcome the laborious currently used manual processes and increase in-depth switch junction analysis, we applied a protocol using high-throughput sequencing (HTS). As a proof of concept, this was done with data generated in our laboratory using the Ion Torrent technology (Life Technologies, Thermo Fisher Scientific) based on simple library preparation utilizing DNA segments from amplification of switch junctions by nested PCR. For processing of large data sets from HTS, available Blastn and BWA algorithms for sequences analysis are inappropriate, because they are unable to automatically summarize results and structural patterns of junctions. The massive amount of HTS data thus requires a specifically designed tool to process sequencing data into insightful information. A big challenge was to develop a new computational tool that automatically analyzes HTS data, which identifies CSR junctions between two regions of the *IgH* locus and summarizes breakpoints and structure (blunt junctions, junctions with microhomology, or insertions inserted into the junction). During this study, we developed CSReport, a program processing HTS data for automatic analysis of CSR recombination junctions. CSReport is a Python code using Blast⁺ and Biopython module. It runs in ~10 min on a laptop computer for a typical 1 million read library. Our results validate CSReport as a successful and accurate computational tool for supporting analysis of extremely large numbers of sequences (as from our HTS-based protocol for CSR junctions), gaining in-depth assessment, and facilitating manipulation and extraction of valuable information.

Materials and Methods

CH12F3 cell culture and CSR induction in vitro

CH12F3 cells were maintained in RPMI 1640 supplemented with Ultra Glutamine, 10% FCS (Lonza), sodium pyruvate (Lonza), penicillin/streptomycin (Lonza), nonessential amino acids (Lonza), and 2-ME. Cells were stimulated to undergo CSR toward IgA with murine IL-4 (5 ng/ml; PeproTech), human TGF- β 1 (1 ng/ml; R&D Systems), and murine anti-CD40 Ab (200 ng/ml; eBioscience) (8). After 72-h activation, cells were labeled with anti-mouse IgA (IgA-FITC; BD Biosciences-Pharmingen) and anti-mouse IgM-PE-cyanine7 (eBioscience) Abs. IgA-bearing cells, representing the population that had undergone CSR, were sorted by a BD FACS Aria III cell sorter for DNA extraction.

Mice and CSR induction in vitro

C57BL/6J wild-type mice (6–8 wk old) were maintained in a specific opportunistic pathogen-free facility and immunized i.p. with 100 μ g nuclear protein–keyhole limpet hemocyanin (Biosearch Technologies) or 50 μ g nuclear protein–AECM–Ficoll (Biosearch Technologies) supplemented with CFA (Sigma-Aldrich). A boost was performed 3 wk later. After 7 d, animals were sacrificed and splenocytes were recovered. Splenocytes were stimulated in vitro to undergo CSR toward Igy1 with LPS B4 (1 μ g/ml; InvivoGen) and murine IL-4 (20 ng/ml; PeproTech) for 72 h. Experimental procedures were performed in accordance with the ethics committee review board of Limoges University (Limoges, France) and Hospital (Comité Régional d'Éthique sur l'Expérimentation Animale du Limousin, Centre Hospitalier Universitaire Dupuytren, Limoges, France).

Amplification of S μ /S α junctions from CH12F3 cells

Genomic DNA from sorted CH12F3 cells was extracted by the phenol/chloroform/isoamyl alcohol method. S μ /S α junctions were amplified by

nested PCR using Phusion HF polymerase from BioLabs. The first PCR used the following primers: Smum1 (5'-CAG TTG AGG CCA GCA GGT-3') and mC α CH1 reverse 2 (5'-TTC TTG GAC GGC GTT AGA GT-3'), and the following protocol: 94°C for 5 min; 35 cycles of 94°C for 1 min, 60°C for 1 min, and 72°C for 1:30 min; and 72°C for 7 min. The second PCR was performed with Smum2 (5'-CAG GTC GGC TGG ACT AAC TC-3') and mC α CH1 reverse 1 (5'-ATG TTG CAC GGA ACA TTT CA-3') under these conditions: 94°C for 5 min; 35 cycles of 94°C for 1 min, 62°C for 1 min, and 72°C for 1 min; and 72°C for 7 min. PCRs were done in triplicate.

Amplification of S μ /S γ 1 junctions from mouse splenocytes

CSR junctions were amplified from murine splenocyte genomic DNA by nested PCR. First PCR primer pairs were as follows: S μ nest1 forward (5'-AGA GAC CTG CAG TTG AGG CC-3') and sg1D reverse (5'-AGG ATG TCC ACC CTC ACC CAG GC-3'), and touchdown conditions as follows: 98°C for 30 s; 2 cycles of 98°C for 10 s, 69°C for 30 s, and 72°C for 90 s; 2 cycles of 98°C for 10 s, 67°C for 30 s, and 72°C for 90 s; 2 cycles of 98°C for 10 s, 65°C for 30 s, and 72°C for 90 s; and 25 cycles of 98°C for 10 s, 60°C for 30 s, 72°C for 90 s, and a final step at 72°C for 7 min. Primers for the second PCR were as follows: S μ Nest forward 2 (5'-CCA GCC ACA GTA ATG ACC CAG-3') and DownSg1 reverse (5'-TTG TTA TCC CCC ATC CTG TCA CCT-3'), according to the following protocol: 98°C for 30 s, and 35 cycles of 98°C for 10 s, 65°C for 30 s, 72°C for 90 s, and 72°C for 7 min. PCRs were done in triplicate for each sample.

Amplification of S μ /S α junctions from human PBMCs

CSR junctions were amplified as previously (9). Amplified CSR junctions from DNA of PBMCs from four healthy individuals were equally mixed to prepare sequencing library. Study was approved by the ethics committee (N°: CPP12-012/2012-A00630-43, Comité de Protection des Personnes du Sud-Ouest et Outre-Mer IV).

Library preparation for ion proton sequencing

Libraries of 200-bp read length were prepared from 100 ng equally mixed triplicate PCR products as a starting material. A first enzymatic fragmentation was performed with Ion Xpress Plus Fragment Library kit (Life Technologies). Fragmented products were then linked to barcodes to identify each library. Ligated DNA fragments were migrated on E-Gel SizeSelect 2% agarose gels, and fragments of desired length were collected and amplified with cyclor parameters, as follows: 95°C for 5 min, and 5 cycles of 95°C for 15 s, 58°C for 15 s, and 70°C for 1 min. Library concentrations were determined by Qubit DsDNA high sensitivity (Invitrogen), and an equimolar mixture was prepared. Libraries were run on an Ion PI v3 chip on the Ion Proton sequencer (Life Technologies).

Sequences are available on European Molecular Biology Laboratory–European Bioinformatics Institute European Nucleotide Archive (<http://www.ebi.ac.uk/ena>; study accession number: ERP019632).

Pipeline HTS data analysis

Raw sequencing data consisted of fastq files containing adapter-trimmed reads and were obtained from the sequencer's processing pipeline. The algorithm of CSReport aims at interpreting each read as an assembly of a segment from the donor S region S μ and segment from an acceptor S region. First, these reads were used as query sequences in a Blast search (Blastn) (10) with ad hoc options (-dust no-, parameter adapted for repeated regions, and -culling_limit 1- restricting the junction mapping to segments with highest score). Reference sequences were slices from the murine C57BL/6 chromosome 12 sequence (gi|372099098 GRCm38.p3) and from the *IgH* locus sequence from *Homo sapiens* chromosome 14 reference (NC_000014.9). Only reads that map both regions, that is, a first hit in donor S region S μ and a second hit in the target S region (user defined), with a minimum of 30 bp for each mapping, were kept for further analysis. A first junction structure evaluation was performed based on segmentation given by Blastn algorithm. The junction structure evaluation consists in determining the number of bases in the overlap between the 3' end of the first segment and the 5' end of the second segment. If this count is strictly positive, the junction is identified as microhomology. If no overlap is found, the junction is identified as either blunt or insertion if some bases are not mapped between the segments' ends. When a long (>10-bp) segment in the middle of a read was not mapped, a second Blast search was performed for the inserted segment (Blastn-short). This rescued some reads that may have originated from multiple breaks in S regions (complex junctions), as already reported (11, 12). When a microhomology (>2 bp) was found, realignment was performed to ensure that reported microhomologies did not contain any mismatches (as the first Blastn algorithm may have introduced mismatches to extend alignment); this

alignment against each germline segment was performed using pairwise 2 from Biopython module (Python implementation of Needleman–Wunsch algorithm). Results of segmentation and junction structure for each filtered read were then reported in a tabular text file. In a subsequent step, these results were processed by statistical summarization procedures that may be selected by user at time of execution as they depend on the validated definition of an individual unique junction: by default, a junction is uniquely defined by one structure/size and two breakpoint positions, and all reads sharing these features are clustered into a single junction's identification. Clustering based on breakpoint proximity can also be used to cluster junctions with very close breakpoints. These steps resulted in a consolidated set of identified junctions described by structure and breakpoint positions. This set was then used to report the structural profile and breakpoint distributions. CSReport workflow is described on the diagram in Supplemental Fig. 1. CSReport as well as other postprocessing scripts are available upon request as a Python code. To use CSReport on other species, one just has to provide an *IgH* reference sequence and an annotation file for S region definition.

Results

Assessment of CSReport using a simulated data set of switch junctions

Use of simulated data sets is a common method to test sensitivity and resolution of computational tools (13) as they compare resulting outputs with expected outcomes. First, CSReport was used to analyze a computer-generated sequence file (Fasta) of 5000 randomly chosen concatenations of germline segments of S_{μ} and S_{α} (each 50–100 bp in length). By generation principle, all sequences are thus blunt junctions

and each of their breakpoints is known. Fig. 1A shows structural profiles of such junctions in terms of either junctions with insertions (more simply insertions), blunt junctions, or junctions with microhomologies (more simply microhomologies). As expected, blunt junctions are the most frequent reported structure (~50% of sequences). Microhomologies are also found at significant frequencies. As already emphasized by previous studies (4), junction analysis is inherently biased to microhomologies as such structures can arise by chance, and this statistical effect (which can be analytically characterized for random sequences) is further increased when recombined regions share extensive homology. If breakpoints are located in short homologous regions, each alignment will extend further than the actual breaks and a microhomology will be found independently of a microhomology-mediated repair mechanism. Finally, it has to be stressed that about only two thirds of junctions (3278 of 5000) were reported by CSReport: a manual curation of some unidentified junctions showed that the homology between S_{μ} and S_{α} was again responsible for this loss of sensitivity as the initial Blastn search failed to properly address each segment to its correct S region.

Accuracy of junction analysis can also be evaluated using a simulated data set of S_{μ} - S_{α} sequences. As shown in Fig. 1B and 1C, segment assignment using CSReport was highly accurate as 90 and 97% of break positions were correctly identified within 10 bp in S_{μ} and S_{α} , respectively. Insets showed a slight dispersion of break positions toward longer alignments for the same reasons

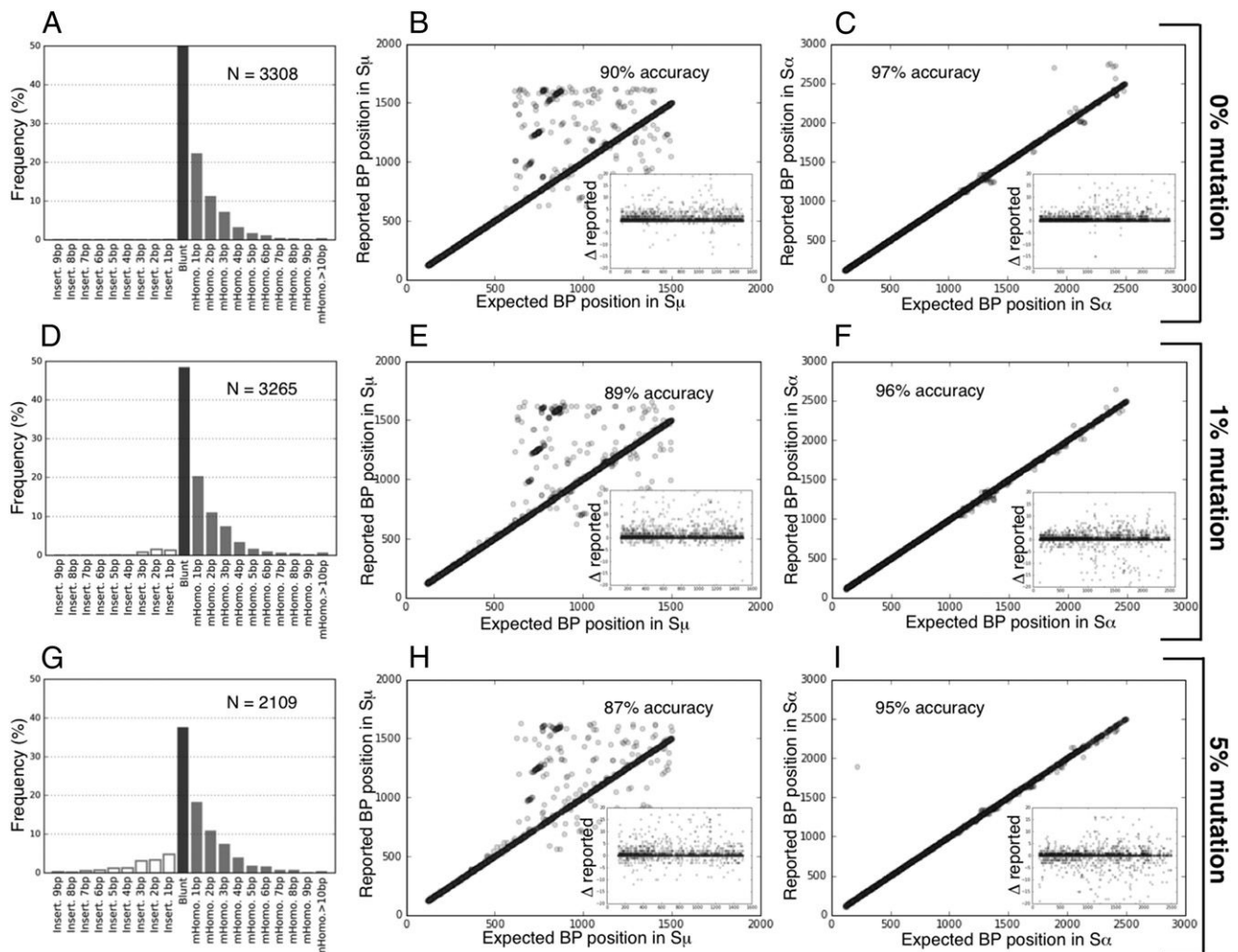


FIGURE 1. CSReport accurately identifies switch junctions in simulated S_{μ} - S_{α} sequences. Structural profiles and reported breakpoint positions are obtained from simulated S_{μ} - S_{α} blunt junctions, randomly generated from concatenation of S_{μ} and S_{α} segments: (A–C) without mutations, (D–F) with 1% random mutations, and (G–I) with 5% random mutations.

as the observed levels of statistical microhomologies. Therefore, the accuracy of junction assignment by CSReport has to be evaluated with some tolerance in break positions (in this study, 10 bp) as this bias is only related to the extent of homology between both S regions and will be inherent to any alignment-based method.

As mutations frequently occur close to breakpoints of switch junctions, CSReport had to be tested in mutated sequences. Simulated mutated data sets were generated with a uniform substitution rate of 1 and 5% and analyzed with CSReport. Fig. 1 shows sensitivity and accuracy results for mutated junctions. As expected, the strongest effect is obtained with 5% mutations. Compared with unmutated sequences, fewer junctions were identified (2109 versus 3278). The

structural profile (Fig. 1G) not only displays the previously discussed bias to microhomologies but also shows insertions at low frequencies. Breakpoint assignment accuracy was also slightly reduced but remained at high levels so that segments were correctly identified for most junctions (Fig. 1H, 1I). Nevertheless, 1% mutation does not seem to modify results compared with no mutations (Fig. 1D–F).

Statistics of $S\mu$ -S α CSR junctions in CH12F3 cells

CSReport was used to analyze switch junctions from in vitro CSR-induced CH12F3 cells. Results from three independent samples are reported in Fig. 2. To document bottlenecks in the pipeline and influence of filter parameters, counts of reads/junctions are shown in Fig. 2A. For each sample, filtration of reads that map to a valid

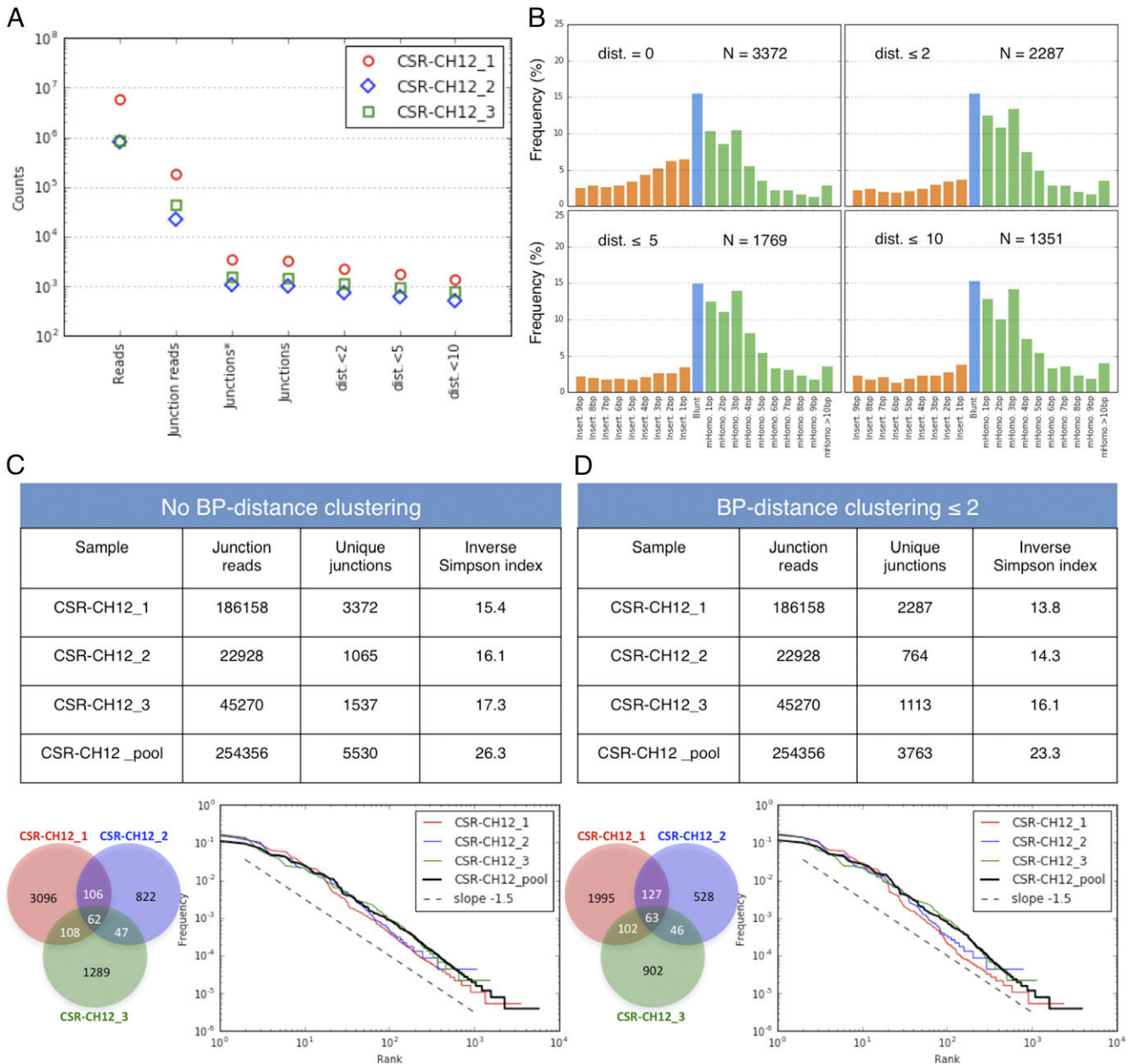


FIGURE 2. Quantitative assessment of switch junctions in CSR-induced CH12F3 cells. **(A)** Overview of junction counts for three independent samples of CH12F3 cells. Reads: sequences in raw fastq file (Ion Proton output). Junction reads: sequences that can be segmented in a strand-aware consistent switch junction. Junctions*: uniquely defined by two breakpoint positions and one structure. Junctions: same as before, and >10-bp insertions are excluded. dist < x: junctions with close breakpoint positions are counted only once. **(B)** Structural profile of sample CSR-CH12_1 for four different values of the breakpoint-distance clusterization parameter (similar trends were observed for the other samples). **(C and D)** Statistics of junction diversity across samples for junctions obtained without breakpoint-distance clusterization (C) or with clusterization of breakpoint distance ≤ 2. Venn diagram comparing the three independent samples. Rank–frequency curves of unique junctions are compared with a generalized Zipf law $f \approx k^{-1.5}$.

CSR junction was highly selective as only 3–5% of input sequences were kept at this step. CSR quantification in a sample relied on determining the number of independent switch recombination events and therefore depended on the definition of a unique junction. A logical choice was to define a junction by its structure (insertion, blunt, or microhomology/size) and the positions of both breakpoints. When counting junctions, 2–5% of junction reads were found to be independent, indicating that most junctions were sequenced multiple times. A reasonable choice was to exclude junctions with large inserts (>10 bp) as they are only partially assembled and represent a small fraction of identified junctions. A further step of summarization was done by clustering identified junctions with very close breakpoints on both S regions. This was parameterized by the maximal distance in base pairs between break positions for two junctions to be clustered. From Fig. 2B, counts of independent junctions after clustering were seen to significantly depend on the distance parameter, whereas the structural profile did not depend on this parameter for non-null values. It thus indicates that distance clustering is efficient in identifying junctions that differ by only a few base pairs in breakpoint positions. A manual curation of these examples suggests this arises when an insertion is present right at the junction site. It is likely that this insertion is an unfortunate sequencing

error (indels are by far the most frequent errors in the Ion Proton system), and these reads should not be considered as single junctions. Increasing the distance clustering parameter further than 2 bp would only reduce the number of independent junctions and seems an overcorrection of this effect as the structure profiles remain unchanged. Detailed investigation of clustering effects is shown in Fig. 2C and 2D in terms of junction counts and diversity. For all three samples, counts were indeed reduced after clustering, but it has to be noted from the Venn diagrams that only nonredundant junctions were significantly removed. Also, sample diversity was slightly reduced after clustering. All these results on three independent samples clearly demonstrated homogeneity (as shown on rank–frequency graphs) so that the pooled data set provided a deeper insight into CSR junctions under these conditions, with greater statistical power.

Summarization of 3763 independent junctions provided descriptions of $S\mu$ - $S\alpha$ CSR sequences at an unprecedented level. The distribution of breakpoints along S regions thus gave indications of motif targeting during the generation of DSB. Fig. 3 indicates colocalization of breakpoints, both in $S\mu$ - $S\alpha$, with motifs AGCT or GGG. The density of breakpoints shares local maxima with the density of one of these two typically abundant switch motifs. The clear correlation observed in this study indicates that breaks will

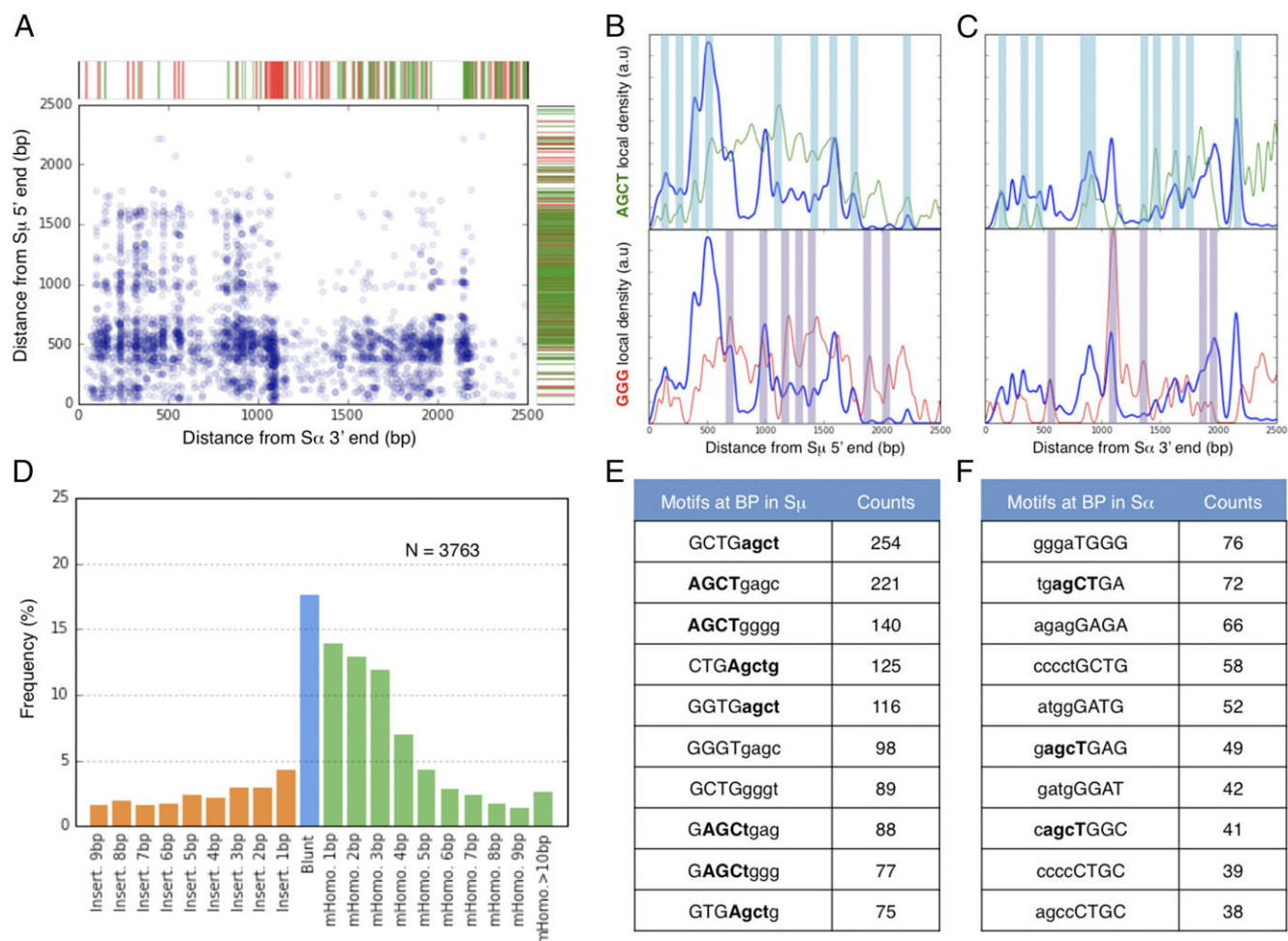


FIGURE 3. CSReport provides valuable DNA break/repair information on large data sets. Structural and positional results obtained from analysis of $S\mu$ - $S\alpha$ junctions in CSR-induced CH12F3 cells (pool of three independent samples). **(A)** Positions of junction breakpoint expressed in terms of distance from S region ends. Each point in the graph is a unique junction defined by 2 bp positions (in $S\mu$ and in $S\alpha$). Positions of palindromic AGCT motifs (green) and GGG motifs (red) are reported on right ($S\mu$) and top sides ($S\alpha$). **(B and C)** Estimation of local density of AGCT motifs (green), GGG motifs (red), and breakpoint along $S\mu$ (B) and $S\alpha$ (C) regions. **(D)** Structural profile of $S\mu$ - $S\alpha$ junctions. **(E and F)** Most frequent 8-nt motifs at breakpoint in $S\mu$ (E) and $S\alpha$ (F). Excised DNA at junctions is shown in lowercase; AGCT motifs are emphasized in bold.

be preferentially generated at local maxima in motif density, regardless of the average motif density of the subregion. As CSReport ensures 1-bp resolution, a more precise characterization of the break mechanism specificity can be obtained by looking at the most frequent nucleotide motif at breakpoints from both regions. Again, both AGCT and GGG were found to account for a large proportion of junction motifs, in accordance with the commonly accepted activation induced-cytidine deaminase-induced break mechanism. Finally, statistics about junction structures (Fig. 3D) can be obtained on a large number of independent junctions. Despite previously discussed bias toward insertions (due to mutations) and microhomologies (by chance, due to region homology), blunt junctions were found to be the most frequent structures. Microhomologies of 1–3 bp were also found at significant levels as the results of shared motifs between S_{μ} and S_{α} . This profile is therefore characteristic of NHEJ-resolved DSB. As already emphasized in a thorough review of CSR junction studies (4), detailed structural information is critical to infer DNA repair pathways. S_{μ} - S_{α} switch junctions are known to be noticeably biased to statistical microhomologies, but studies have been limited to a few manually curated junction sequences. In this work, our experimental HTS-based protocol and our computational analysis with CSReport give valuable detailed and accurate information on a large set of switch junctions for better elucidation of recombination mechanisms.

Statistics of CSR S_{μ} - $S_{\gamma}1$ junctions in resting versus stimulated mouse splenocytes

The comparative analysis of CSR S_{μ} - $S_{\gamma}1$ junctions in resting versus stimulated mouse splenocytes was performed using CSReport, and results are shown in Fig. 4. As expected, in vitro LPS plus IL-4 stimulation of splenocytes efficiently induced CSR as a greater number of independent S_{μ} - $S_{\gamma}1$ junctions was found ($n = 9428$) compared with those before stimulation ($n = 674$). Interestingly, the structural profile was also modified as blunt junctions were reported in a higher proportion of stimulated B cells (27.5%) than in resting B cells (15.9%). Whereas microhomology frequencies exhibited a smooth decrease as a function of microhomology size in stimulated cells, microhomologies of 1–4 bp were found approximately at the same amounts as blunt junctions in resting cells.

Quantitative and qualitative differences can also be documented in terms of breakpoint positions and distinctive motifs at junctions. Fig. 4B shows how possible breakpoint positions were used upon stimulation: breaks resulting in valid junctions span over 3 kb in $S_{\gamma}1$ in stimulated cells, whereas they appear more localized to the 3' end of this region in resting splenocytes. In contrast, distribution of break positions in S_{μ} was similar in both conditions (spanning over the first 1.6 kb). Again, local density maxima of breakpoints in the S regions colocalized with local density maxima of palindromic AGCT and GGG motifs, as highlighted by the prevalence of such



FIGURE 4. CSReport identifies switch junctions in mouse splenocytes and human PBMCs. (A) Structural profiles, (B) distribution of breakpoints, and (C) most frequent 8-nt motifs at breakpoint. Left and middle figures are obtained from S_{μ} - $S_{\gamma}1$ junctions, respectively, in resting and stimulated mouse splenocytes. Right figures are obtained from S_{μ} - S_{α} junctions in human PBMCs.

patterns in the distinctive break sequences (Fig. 4C). Again, S_{μ} breaks exhibited more similarity in motif usage when stimulated cells were compared with resting cells.

Discussion

Sequence analysis of switch junctions from B cells reveals valuable information on the mechanism and effectiveness of CSR, and this is done almost routinely during studies of B cell functions in a variety of mouse models and clinical research. We aimed to study CSR junctions using a protocol based on HTS (Ion Torrent technology and Ion Proton sequencer). This technique challenges currently used methods already used due to the extremely large amount of collected data. To address the lack of automated tools for the analysis of large data sets of switch junction sequences, we designed CSReport, a software that ensures assessment of switch recombination junctions from Fastq or Fasta files. Typically, CSReport analyzes millions of reads that represent a few thousand single CSR junctions (Fig. 2A). In commonly used methods for CSR junction study, bias in junction mapping and junction structural characterization can arise from the chosen alignment algorithm as well as from the author interpretation. Blastn remains the more appropriate mapping software in regard to the BWA aligner, mostly used for HTS data analysis that we excluded because results obtained with BWA were very far from manual analysis. Blastn parameters were set up, but can be modified by the user, and we used the `-no dust` option appropriate to the analysis of repeated sequences as S regions and the `-culling limit-1` returning the highest quality mapping result. This implies that the correct reference sequence is used, which is particularly important for different murine genetic backgrounds. To evaluate errors in junction analysis induced by CSReport, we subjected it to analysis of artificially designed blunt-ended S_{μ} - S_{α} junctions with known breakpoints (Fig. 1). Obtained results showed that CSReport interpretation induced a shift of blunt-ended junctions toward junctions using microhomologies (Fig. 1A) in the same way as a manual analysis, whatever the used aligner. In addition, mutations in proximity to the junction site induced false detection of junctions with insertions (Fig. 1D–G). These effects are most likely to be present in previous methods of junction analysis as well. As HTS data are more prone to sequencing errors, it appeared that clustering of unique junctions can lower the rate of abnormal detection of junctions with insertions (Fig. 2A, 2B). Clustering is controlled by the breakpoint-distance parameter (maximal distance of breakpoint of two individual junctions for them to be clustered). Based on our results, we recommend using CSReport with a breakpoint distance ≤ 2 bp. Increasing further this parameter will only result in fewer reported junctions. It is of note that this 2-bp clustering does not significantly affect diversity of the samples (Fig. 2C, 2D), as evaluated by Inverse Simpson Index, which can be only determined with large amounts of data. More detailed evaluation of sample clonality is provided by rank-frequency curves that also require substantial depth. After accuracy assessment of CSR junction identification by CSReport, we submitted results from HTS of five libraries made using amplified CSR junctions from DNA of CSR-stimulated CH12F3 cells (Fig. 3) and normal mouse splenocytes before and after 3 d of IL-4 plus LPS in vitro stimulation (Fig. 4). CSReport permits the automatic determination of breakpoint positions along the reference sequences, structural features of CSR junctions, and the most frequent DNA motifs at the junction breakpoints. As expected, S_{μ} - $S_{\gamma 1}$ junction count was greatly increased after splenocyte stimulation, and CSReport revealed structural features for all tested samples with 25–50% of blunt-end junctions, whereas

junctions with microhomologies longer than 4 bp were quite rare, according to results in the literature (4).

Finally, we propose CSReport as an automated computational tool able to support analysis of large data sets from CSR junction HTS. Widespread use of CSReport will normalize studies analyzing switch junctions as it provides a single pipeline providing breakpoint positions and junction structure (blunt-end junctions or junctions with insertion or microhomology). CSReport conforms to the definition of microhomology as perfectly matched nucleotides shared by donor and acceptor junction segments without any imperfect homology or mismatch, as previously described (14). CSReport also includes specific identification of complex sequential CSR (S_{μ} - S_x - S_y junctions) by realigning long insertions (>10 bp) that may not have been mapped during the first alignment step. In addition, the use of CSReport will reduce the bias in the delineation of CSR junction characteristics due to operator interpretation. CSReport is able to support Fastq and Fasta files and is meant to process data sets regardless of the sequencing technology used to produce them. It is of note that read size is not a limitation criterion, as in Ig or TCR repertoire analyses (15, 16), which need ~ 500 -bp length sequencing, because CSReport only keeps for analysis those reads that map on donor and acceptor regions for >30 bp. Nevertheless, large fragments could certainly improve capture of junctions containing reads.

Another interesting point is that CSReport permits CSR junction study regardless of the species, as demonstrated by S_{μ} - $S_{\alpha 1}$ junction analysis done on human PBMCs (Fig. 4). Finally, in addition to the count and characterization of CSR junctions, CSReport presents results in the form of clear and easily manipulated tabular files of valuable information. CSReport facilitates and accelerates the study of CSR junctions. It also harmonizes analysis and interpretation.

Acknowledgments

We thank the GéoLim platform and the Cytométrie, Imagerie et Microscopie platform (Génomique Environnement Immunité Santé et Thérapeutique Institut, Université de Limoges, Centre de Biologie et de Recherche en Santé, Centre Hospitalier Universitaire de Limoges) for sequencing and cell sorting, and S. Desforges, B. Remerand, and D. Champagnol for help with animal care.

Disclosures

The authors have no financial conflicts of interest.

References

- 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.
- Teng, G., and F. N. Papavasiliou. 2007. Immunoglobulin somatic hypermutation. *Annu. Rev. Genet.* 41: 107–120.
- Revy, P., T. Muto, Y. Levy, F. Geissmann, A. Plebani, O. Sanal, N. Catalan, M. Forveille, R. Dufourcq-Labouesse, A. Gennery, et al. 2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the hyper-IgM syndrome (HIGM2). *Cell* 102: 565–575.
- 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.
- Soulas-Sprauel, P., G. Le Guyader, P. Rivera-Munoz, V. Abramowski, C. Olivier-Martin, C. Goujet-Zalc, P. Charneau, and J.-P. de Villartay. 2007. Role for DNA repair factor XRCC4 in immunoglobulin class switch recombination. *J. Exp. Med.* 204: 1717–1727.
- 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.
- Pan, Q., H. Rabbani, F. C. Mills, E. Severinson, and L. Hammarström. 1997. Allotype-associated variation in the human $\gamma 3$ switch region as a basis for differences in IgG3 production. *J. Immunol.* 158: 5849–5859.
- Nakamura, M., S. Kondo, M. Sugai, M. Nazarea, S. Imamura, and T. Honjo. 1996. High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. *Int. Immunol.* 8: 193–201.

9. Pan, Q., C. Petit-Frère, S. Dai, P. Huang, H. C. Morton, P. Brandtzaeg, and L. Hammarström. 2001. Regulation of switching and production of IgA in human B cells in donors with duplicated alpha1 genes. *Eur. J. Immunol.* 31: 3622–3630.
10. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403–410.
11. 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. [Published erratum appears in 2006 *J. Exp. Med.* 203: 251.] *J. Exp. Med.* 203: 99–110.
12. Reina-San-Martin, B., J. Chen, A. Nussenzweig, and M. C. Nussenzweig. 2007. Enhanced intra-switch region recombination during immunoglobulin class switch recombination in 53BP1^{-/-} B cells. *Eur. J. Immunol.* 37: 235–239.
13. Safonova, Y., A. Lapidus, and J. Lill. 2015. IgSimulator: a versatile immunosequencing simulator. *Bioinformatics* 31: 3213–3215.
14. Pan, Q., C. Petit-Frère, A. Lähdesmäki, H. Gregorek, K. H. Chrzanowska, and L. Hammarström. 2002. Alternative end joining during switch recombination in patients with ataxia-telangiectasia. *Eur. J. Immunol.* 32: 1300–1308.
15. IJspeert, H., P. A. van Schouwenburg, D. van Zessen, I. Pico-Knijnenburg, A. P. Stubbs, and M. van der Burg. 2017. Antigen Receptor Galaxy: a user-friendly, web-based tool for analysis and visualization of T and B cell receptor repertoire data. *J. Immunol.* 198: 4156–4165.
16. Langerak, A. W., M. Brüggemann, F. Davi, N. Darzentas, D. Gonzalez, G. Cazzaniga, V. Giudicelli, M.-P. Lefranc, M. Giraud, E. A. Macintyre, et al. 2017. High-throughput immunogenetics for clinical and research applications in immunohematology: potential and challenges. *J. Immunol.* 198: 3765–3774.

Figure S1

A

Fasta or Fastq file from any sequencing platform
(barcode and adapter-trimmed)

Seq.fasta

CSReport_run(Seq,Ref,Region1,Region2)

Seq_ReadsJunctions.csv

Read	Segmentation	Start1 (read)	End1 (read)	Start1 (ref)	End1 (ref)	Start2 (read)	End2 (read)	Start2 (ref)	End2 (ref)	Junction type	Junction size
QLFG:01337:11602_rc	Salpha-Smu	1	34	52671	52704	36	73	217665	217702	Insert	1
QLFG:01338:11636_rc	Salpha-Smu	1	209	52417	52622	210	256	217470	217516	Blunt	0
QLFG:01342:11685	Salpha-Smu	1	165	52880	53044	162	223	217429	217490	mHomo	4

CSReport_summary(Seq,cluster=2)

Seq_Junctions.csv

Junction	Count
54424-217654-Insert-1	6
53085-217387-Blunt-0	2
52383-217430-Blunt-0	1
53322-217773-mHomo-4	1

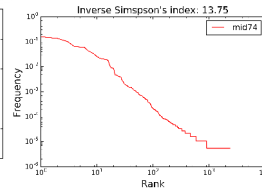
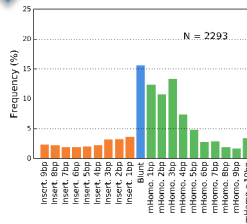
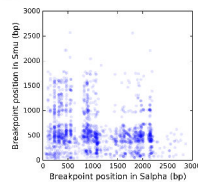
CSReport post-processing

Motifs at BP in S_{μ}

GCTGagct	169
AGCTgagc	161
AGCTgggg	113
GGTgagct	102
CTGAgctg	97

Motifs at BP in S_{α}

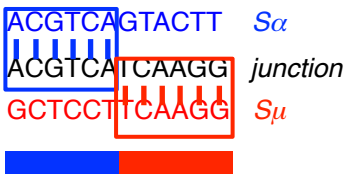
gggaTGGG	79
ccctGCTG	69
tgagCTGA	69
ggatGGGA	56
ccacTGCC	48



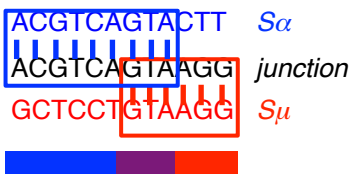
CSReport WORKFLOW

B

Blunt



Micro-homology (3bp)



Insertion (2bp)

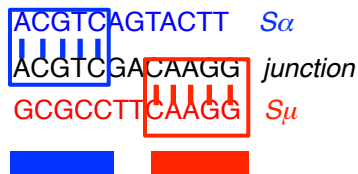


Figure S1. CSReport identifies swith junctions from raw sequencing reads and provides graphical results for direct interpretation. (A) Workflow diagram of CSReport. Input files are sequencing reads in fasta or fastq format. The first command (CSReport_run) calls Blastn and assembles junctions from realignment of segments'ends (individual structure and BP). The second command (CSReport_summary) identifies single junctions from the junction reads, accordingly to the user-defined parameter of distance clustering. Other CSReport functions produces graphical output to represent structural profile, BP distribution, junctional diversity or involved nucleotidic motifs. (B) Definition of blunt junction (left), junction with micro-homology (middle), or junction with insertion (right). CSReport conforms to those definitions

Annexe 2: (Article)

G-quadruplex DNA targeting alters class-switch recombination in B cells and attenuates allergic inflammation.

Zeinab Dalloul, Pauline Chenuet, **Iman Dalloul**, François Boyer, Jean-Claude Aldigier, Brice Laffleur, Yolla El Makhour, Bernhard Ryffel, Valerie F. J. Quesniaux, Dieudonnee Togbé, Jean-Louis Mergny, Jeanne Cook-Moreau, Michel Cogne.

Article accepté à Journal of Allergy and Clinical Immunology JACI en juin 2018.

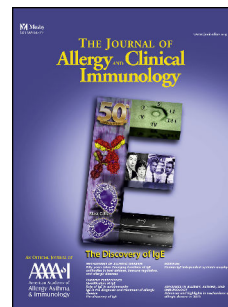
Cette étude vise à tester l'effet d'une drogue pharmacologique (RHPS4) connue comme stabilisateur des structures « G-quadruplex » (et précédemment testé en tant que potentiel agent anti-cancéreux du fait de sa capacité à inhiber certains promoteurs d'oncogènes) sur le mécanisme de la CSR. Nous avons montré que le RHPS4 réduit de manière significative la CSR mais aussi la sécrétion des différentes classes d'Ig par les cellules B ; il diminue aussi la sécrétion d'IgE et d'IgG1 spécifique de l'ovalbumine après une immunisation in-vivo. Ce traitement a aussi montré un rôle important dans l'atténuation des signes de l'asthme allergiques, indiquant de potentielles pistes thérapeutiques pour des inhibiteurs du switch (par exemple pour traiter l'asthme en ciblant la commutation de classe vers IgE et IgG1).

Dans cette publication, j'ai participé dans les manip : Prélèvement des splénocytes et stimulation in-vitro, extraction d'ARN et RT-qPCR, PCR et préparation des bibliothèques pour le séquençage Ion Proton NGS. J'ai participé aussi à faire les figures de cette publication.

Accepted Manuscript

G-quadruplex DNA targeting alters class switch recombination in B cells and attenuates allergic inflammation

Zeinab Dalloul, MSc, Pauline Chenuet, MSc, Iman Dalloul, MSc, François Boyer, PhD, Jean-Claude Aldigier, MD, PhD, Brice Laffleur, PhD, Yolla El Makhour, PhD, Bernhard Ryffel, MD, PhD, Valerie FJ. Quesniaux, PhD, Dieudonnée Togbé, PhD, Jean-Louis Mergny, PhD, Jeanne Cook-Moreau, PhD, Michel Cogné, MD, PhD.



PII: S0091-6749(18)30895-9

DOI: [10.1016/j.jaci.2018.06.011](https://doi.org/10.1016/j.jaci.2018.06.011)

Reference: YMAI 13486

To appear in: *Journal of Allergy and Clinical Immunology*

Received Date: 31 January 2018

Revised Date: 30 May 2018

Accepted Date: 12 June 2018

Please cite this article as: Dalloul Z, Chenuet P, Dalloul I, Boyer F, Aldigier J-C, Laffleur B, El Makhour Y, Ryffel B, Quesniaux VF, Togbé D, Mergny J-L, Cook-Moreau J, Cogné M, G-quadruplex DNA targeting alters class switch recombination in B cells and attenuates allergic inflammation, *Journal of Allergy and Clinical Immunology* (2018), doi: 10.1016/j.jaci.2018.06.011.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

G-quadruplex DNA targeting alters class switch recombination in B cells and attenuates allergic inflammation

Zeinab Dalloul^a , MSc Pauline Chenuet^b, MSc, Iman Dalloul^a, MSc, François Boyer^a, PhD, Jean-Claude Aldigier^a, MD, PhD, Brice Laffleur^a, PhD, Yolla El Makhour^c, PhD, Bernhard Ryffel^b, MD, PhD, Valerie FJ Quesniaux^b, PhD, Dieudonnée Togbé^b, PhD, Jean-Louis Mergny^{d, e}, PhD, Jeanne Cook-Moreau^a , PhD & Michel Cogné^{a,f}, MD, PhD.

^a Limoges University, CNRS UMR 7276, INSERM U1162, :Control of the B cell Response & Lymphoproliferation, 2 rue du Dr. Marcland, 87025 Limoges, France

^b INEM - UMR7355, CNRS, 3b rue de la Férollerie, F-45071 Orléans - Cedex 2. / ArtImmune SAS, 13 Avenue Buffon, 45100, Orléans

^c Immunology Unit MICSU and Lebanese University, Faculty of sciences, Beirut, Lebanon.

^d Bordeaux University, INSERM U1212, UMR CNRS 5320, ARNA, 33600 Pessac, France.

^e Institute of Biophysics of the CAS, v.v.i., Královopolská 135, 612 65 Brno, Czech Republic.

^f Address correspondence to Michel Cogné, MD, PhD, 2 rue du Dr. Marcland, 87025 Limoges, France. Fax # (33) 555 435 897, Phone (33) 519 564 200, mail: cogne@unilim.fr

The authors declare no conflict of interest.

Capsule summary: Targeting immunoglobulin *switch* regions with a G4-DNA binding drug offers a novel immunomodulatory strategy reducing class-switched antibody secretion.

Keywords: Allergic asthma model ; immunoglobulin class switch ; B lymphocytes ; G-quadruplex ; immunopathology

Abbreviations: AID, Activation-induced deaminase. BALF, Bronchoalveolar lavage fluid.

CSR, Class-switch recombination. G4, G-quadruplex. RHPS4, 3,11-difluoro-6,8,13-trimethyl-

8*H*-quino[4,3,2-*kl*]acridinium methosulfate.

ACCEPTED MANUSCRIPT

G-quadruplex DNA targeting alters class switch recombination in B cells and attenuates allergic inflammation

To the Editor

G-quadruplex (G4) DNA abundance in immunoglobulin (Ig) switch (S) regions (see Fig E1 and the Methods section in this article's Online Repository at www.jacionline.org) might promote transcriptional pausing and occurrence of DNA nicks¹. *In vitro*, such structured DNA favors oligomerization and activity of Activation-Induced Deaminase (AID)². Orientation-dependent CSR optimally occurs with G-rich DNA on the non-template strand³. However, AT-rich *Xenopus laevis* S-regions support CSR in an orientation-independent manner and it was rather proposed that palindromes forming stem loops are mandatory for CSR, the role of G4-DNA remaining controversial⁴.

Pharmacological G4-stabilizing agents are potential anticancer drugs with telomerase inhibitory activity and potential to slow expression of some oncogene promoters⁵. Their effect on CSR is unknown. We hypothesized that if G4-DNA promotes AID recruitment onto S-regions, G4-stabilizing agents should then affect CSR. We evaluated non-toxic doses of the prototypic G4-ligand, 3,11-difluoro-6,8,13-trimethyl-8*H*-quino[4,3,2-*kl*]acridinium methosulfate (RHPS4) both *in vitro* on cultured B-lymphocytes and *in vivo* in mice, with regards to CSR and B-cell terminal differentiation.

While 5 to 50 μ M RHPS4 can target cancer cells, we used doses in the 0.1 to 2 μ M range *in vitro*, and 10 mg/kg/day for *in vivo* assays. Such low doses neither decreased B-cell growth in cultures nor triggered apoptosis but somehow impacted plasma cell differentiation (see Fig E2A-C in this article's Online Repository at www.jacionline.org). Splenocytes recovered from mice treated for 9 days with RHPS4 also showed unaltered viability (see Fig E2D in this article's Online Repository at www.jacionline.org). We validated that such RHPS4

concentrations indeed more strongly bound oligonucleotides homologous to S μ rather than control (see Fig E3 in this article's Online Repository at www.jacionline.org).

To evaluate whether RHPS4 modulated class-switching, splenocytes were stimulated *in vitro* by lipopolysaccharide (LPS) +/- cytokines, and followed for cell-surface BCR expression. This yielded a 3-fold reduction of class-switched cells at 0.5 μ M RHPS4 and a 5-fold reduction at 1 μ M (by either specifically monitoring cells switched to IgG1 or globally evaluating IgM^{neg} CD19⁺ class-switched cells (**Fig 1A**)). IgM secretion in supernatants was barely influenced at 0.5-1 μ M RHPS4 (and cut by 1/3 at 2 μ M), while class switched Ig decreased more than ten-fold (*i.e.* IgG2b & IgG3 with LPS alone, IgG2a with LPS + interferon γ , IgG1 & IgE with LPS+ IL-4) (at any dose over 0.5 μ M RHPS4) (**Fig 1B**). While B-cell counts in such cultures did not vary, decreased differentiation into CD138⁺ plasmablasts (cut by about one-third) (**Fig. E2C**) might contribute to globally reduced Ig secretion.

We quantified I γ -C γ and I μ -C μ germline transcripts specific for the "pre-CSR" stage and I μ -C γ "post-CSR" transcripts in LPS/IL-4-stimulated B-cells. Upon RHPS4 treatment, "pre-CSR" transcripts increased (**Fig 1C, left panels**), indicating that IgH locus transcription preceding CSR accessibility was maintained. By contrast, class-switched transcripts decreased (**Fig 1C, right panels**), indicating a blockade of the CSR step itself.

Although AICDA gene transcription was unaffected by RHPS4, chromatin immunoprecipitation (ChIP) experiments confirmed a drastic reduction in AID-recruitment to S μ , S ϵ as well as S γ 1 in stimulated cells (**Fig 1D, E**).

S μ -S γ CSR junctions sequenced by NGS from cells treated with RHPS4 featured a 4-fold decrease in number, with slightly increased proximity of breaks to cytosines of RGYW/WRCY motifs, while the balance of non-homologous-end joining *versus* microhomology-mediated-end joining was unchanged (see Fig E4 in this article's Online Repository at www.jacionline.org).

After primary immunization by ovalbumin of mice receiving RHPS4, we observed that specific anti-OVA unswitched IgM-producing cells increased (as evaluated by ELISPOT), while IgG secreting cell number was cut by half, and serum circulating anti-OVA-specific IgG (and IgE) accordingly decreased (**Fig 2A**). To a lesser extent, serum anti-OVA specific IgM also decreased, suggesting that on a “per-cell” basis, IgM secretion also tended to be lower.

We followed secondary responses in mice booster-immunized at Day 7 and 14 and kept for 25 days under either steroids or RHPS4. This indicated a prolonged decrease in specific IgG and delayed production of IgE (decreased at Day 9 but not Day 25) (**Fig 2B**).

In mice sensitized to OVA, allergic manifestations triggered by intra-nasal immunization were explored under treatment by RHPS4 or steroids. Following the attenuated systemic OVA-specific IgG1 and IgE responses mentioned above, some aspects of airway allergy were alleviated in lungs of mice receiving RHPS4: inflammation scores and cell infiltration decreased in lungs (**Fig 2C, D**). Cell-numbers were reduced for all cell types in broncho-alveolar fluid (BALF), in comparison to mice only receiving ovalbumin (most significantly for macrophages and lymphocytes), and RHPS4 significantly reduced total protein levels in BALF (**Fig 2E**). Most of these biological markers of RHPS4-mediated immunomodulation reached levels similar to those achieved with steroids and were most likely related to the attenuated humoral response, since in parallel, T-cell parameters such as cytokine levels were improved as expected by steroids but unaffected by RHPS4 treatment (see Fig E5 in this article’s Online Repository at www.jacionline.org). Consequently, some but not all functional aspects of airway allergy (airway compliance and lung eosinophil peroxidase activity) improved under RHPS4 treatment (see Fig E6 in this article’s Online Repository at www.jacionline.org).

Inappropriate humoral immune responses involving pro-inflammatory class-switched Igs frequently result in immunopathology, and are then treated with broad anti-inflammatory

strategies. Few compounds currently target B-cell function and none directly target CSR. It would therefore be of considerable interest to identify drugs directly modulating CSR, thus lessening production of the most pro-inflammatory class-switched Ig. CSR in B-cells is modulated by multiple *cis*-regulatory elements, notably 3' IgH enhancers promoting AID recruitment^{6,7}. Convergent data suggest that CSR is favored by G-rich S-region transcripts forming RNA/DNA hybrids on the template strand, exposing single-stranded R-loops on the non-template strand, and differentially recruiting AID to initiate double-strand breaks instead of mutations³. G4-DNA defining “G-loops” on the non-template strand may also help recruit CSR factors⁸. AID binding to G4-containing transcripts, might finally confer *trans* activity to such transcripts⁹.

We now show that low doses of a G4-DNA stabilizing agent documented for anti-cancer activity, directly dampen CSR *in vitro*. Despite broader effects expected *in vivo* due to the presence of G4-DNA in many genes, RHPS4 lowered numbers of switched B-cells after immunization, impacted production of class-switched Ig and attenuated manifestations of allergy in a mouse model of airway sensitization (see graphical abstract as Fig E7 in this article's Online Repository at www.jacionline.org). This report thus shows CSR as a “druggable” process, with G4-stabilizing agents such as RHPS4 both inhibiting CSR and decreasing Ig secretion, both of which might be of interest in auto-immune or immuno-allergic conditions.

Zeinab Dalloul^a, Pauline Chenuet^b, Iman Dalloul^a, François Boyer^a, PhD, Jean-Claude Aldigier^a, MD, PhD, Brice Laffleur^a, PhD, Yolla El Makhour^c, PhD, Bernhard Ryffel^b, MD, PhD, Valerie FJ Quesniaux^b, PhD, Dieudonnée Togbé^b, PhD, Jean-Louis Mergny^d,^e, PhD, Jeanne Cook-Moreau^a, PhD & Michel Cogné^{a,f}, MD, PhD.

^a Limoges University, CNRS UMR 7276, INSERM U1162, :Control of the B cell Response & Lymphoproliferation, 2 rue du Dr. Marcland, 87025 Limoges, France

^b INEM - UMR7355, CNRS, 3b rue de la Férollerie, F-45071 Orléans - Cedex 2. / ArtImmune SAS, 13 Avenue Buffon, 45100, Orléans

^c Immunology Unit MICSU and Lebanese University, Faculty of sciences, Beirut, Lebanon.

^d Bordeaux University, INSERM U1212, UMR CNRS 5320, ARNA, 33600 Pessac, France.

^e Institute of Biophysics of the CAS, v.v.i., Královopolská 135, 612 65 Brno, Czech Republic.

^f Correspondence: cogne@unilim.fr

Acknowledgements: This work was supported by grants from ARC (PGA120150202338), ANR (grant 16-CE15-0019-01), INCa (grant #9363), Ligue Nationale contre le Cancer, European FEDER n°2016-00110366, Biotarget Région Centre. SYMBIT Région Nouvelle Aquitaine (CZ.02.1.01/0.0/0.0/15_003/0000477) and ERDF. ZD was supported by a fellowship from Fondation Française pour la Recherche contre le Myélome et les Gammopathies monoclonales (F.F.R.M.G.). We thank Dr. P. Gearhart for anti-AID antibodies.

REFERENCES

1. Vallur AC, Maizels N. Activities of human exonuclease 1 that promote cleavage of transcribed immunoglobulin switch regions. *Proc Natl Acad Sci USA*. 2008;105:16508–12.
2. Qiao Q, Wang L, Meng F-L, Hwang JK, Alt FW, Wu H. AID Recognizes Structured DNA for Class Switch Recombination. *Mol Cell*. 2017;67:361-373.e4.
3. Shinkura R, Tian M, Smith M, Chua K, Fujiwara Y, Alt FW. The influence of transcriptional orientation on endogenous switch region function. *Nat Immunol*. 2003;4:435–41.
4. Tashiro J, Kinoshita K, Honjo T. Palindromic but not G-rich sequences are targets of class switch recombination. *Int Immunol*. 2001;13:495–505.
5. Balasubramanian S, Hurley LH, Neidle S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat Rev Drug Discov*. 2011;10:261–75.
6. Saintamand A, Rouaud P, Saad F, Rios G, Cogné M, Denizot Y. Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. *Nat Commun*. 2015;6:7084.
7. Le Noir S, Boyer F, Lecardeur S, Brousse M, Oruc Z, Cook-Moreau J, et al. Functional anatomy of the immunoglobulin heavy chain 3' super-enhancer needs not only core enhancer elements but also their unique DNA context. *Nucleic Acids Res*. 2017;45:5829–37.
8. Duquette ML, Handa P, Vincent JA, Taylor AF, Maizels N. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev*. 2004;18:1618–29.
9. Zheng S, Vuong BQ, Vaidyanathan B, Lin J-Y, Huang F-T, Chaudhuri J. Non-coding RNA Generated following Lariat Debranching Mediates Targeting of AID to DNA. *Cell*. 2015;161:762–73.

ACCEPTED MANUSCRIPT

Figure 1. RHPS4 decreases CSR in stimulated primary mouse B cells.

(A) Primary B cells stimulated with LPS/IL-4 were treated with 0, 0.5, and 1 μ M RHPS4, collected after 4-day culture and stained for flow cytometry. Histograms represent mean % \pm SEM for 6 mice. *** P <0.001 compared with untreated cells.

(B) Supernatants from adequately stimulated splenocytes were quantified by ELISA for the production of IgM, IgG2b and IgG3 (*top*), IgG2a, IgG1 and IgE (*bottom*).

(C) Germline C γ 1 and C μ transcripts (*left*) and post-switch (*right*) I μ -C γ 1 and I μ -C ϵ transcripts from LPS+IL4 stimulated B-cells. Histograms represent mean fold change \pm SEM for 6 mice. (D) AICDA expression from LPS-stimulated cells. (E) ChIP experiments showing AID-recruitment to S μ , S ϵ and S γ 1-regions in LPS+IL-4 stimulated cells (*right*).

NS: not significant. * P <0.05, ** P <0.01, *** P <0.001 compared with the control untreated cells.

Figure 2. RHPS4 inhibited class-switching *in vivo* and decreased production of specific antibodies.

(A) Ag-specific IgM or IgG producing cells by ELISPOT in mice after primary immunization (left panels), and levels of Ag-specific Abs (right panels). (B). Ag-specific IgG1 and IgE Abs after 25 days i.p. and i.n. ovalbumin immunization.

(C) In an allergic asthma model after repeated Ag administration, mice receiving RHPS4 displayed less abundant cell infiltration of lung, (D) lower inflammation and mucus scores and (E) lower counts of total cells, lymphocytes, macrophages, eosinophils, and neutrophils, together with lower protein levels, in bronchoalveolar fluid (BALF).

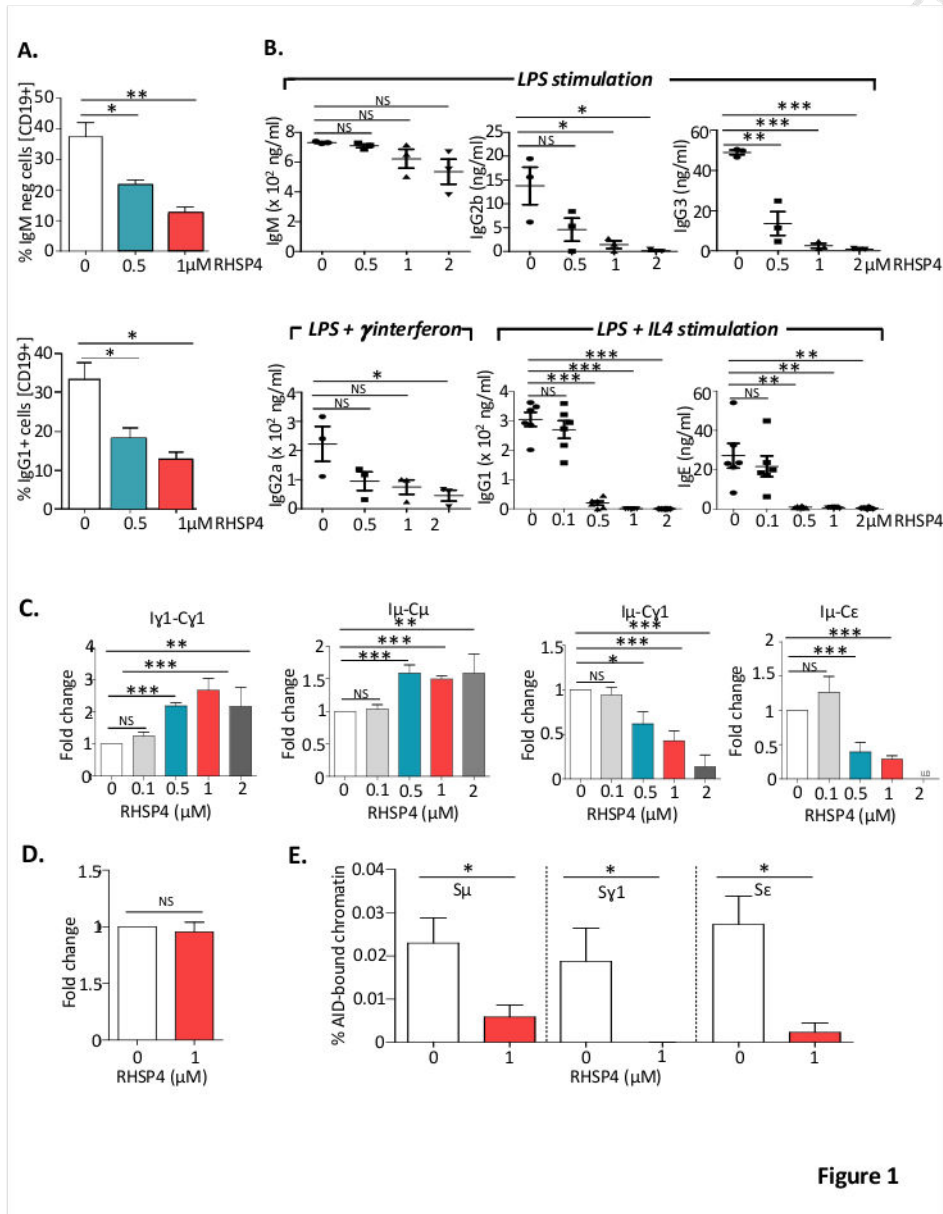
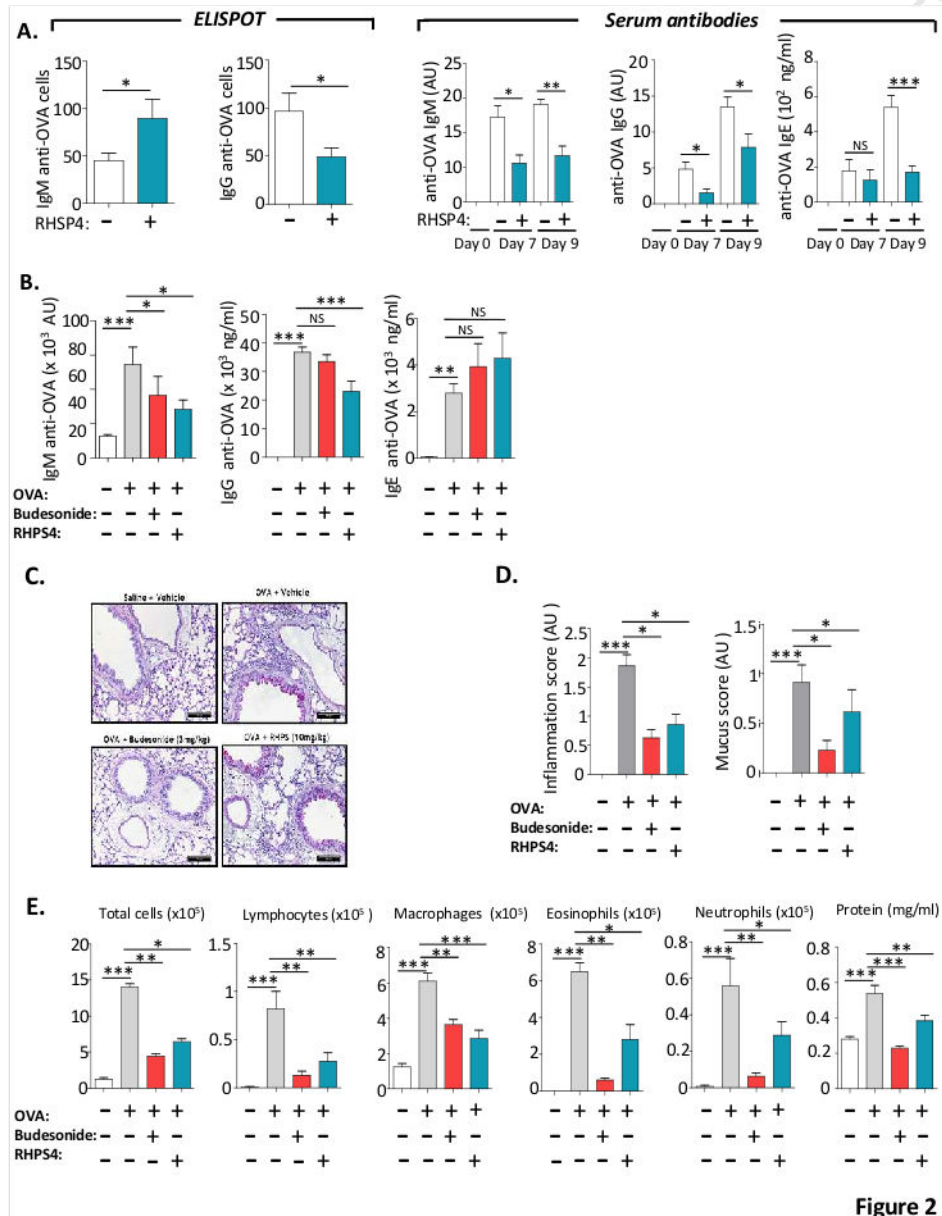


Figure 1



METHODS

Mice

WT mice were used in all experiments and maintained in our animal facilities, at 21–23°C with a 12-h light/dark cycle. Our research was approved by the local ethics committee review board. Three to five month old BALB/c or C57BL/6 mice were used for the *vitro* culture experiments. Six to eight weeks old SV129 and BALB/c mice were used for *in vivo* treatment trials.

Cell cultures

Splenocytes were collected, red blood cells were lysed and CD43⁺ cells were depleted using CD43 microbeads (Miltenyi Biotec). B cells were cultured for 4 days (for ELISA, qPCR and flow cytometry assays) and for 2 days (for ChIP experiments), in RPMI containing 10% FCS with LPS (20 µg/mL) (Invivogen), LPS + IL-4 (40 ng/mL) or LPS + IFN γ (2 ng/mL) (Peprotech) with or without RHSP4 (TOCRIS). RHSP4 was used at concentrations of 0.1, 0.5, 1 and 2 µM depending on the experiment.

Proliferation Assay (MTS)

B cells isolated from 6 mice were seeded into 96-well plates at a density of 1×10^5 per well with a total volume equal to 100 µL, stimulated with LPS + IL-4 and treated with the following concentrations of RHPS4: 0.1, 0.5, 1 and 2 µM. Control cells were stimulated with LPS+IL-4 but not treated. The experiment was stopped at Day 4. For the MTS assay, the CellTiter 96H AQueous One Solution Cell Proliferation Assay kit (Promega) was used following the manufacturer's instructions. Briefly, at 3 h before the desired time point, 20 µl

MTS reagent was added to each well and cells were incubated at 37 °C for 3 h. Absorbance was detected at 492 nm with a Microplate Reader (ThermoScientific ; Morti).

Class-specific ELISA

ELISA was performed on sera and supernatants from *in vitro* stimulated and RHPS4 (0.1, 0.5, 1 and 2 µM) treated and untreated primary B cells for the detection of IgM, IgG1, IgE, IgG2a, IgG2b and IgG3 secretion. Plates were coated with monoclonal antibodies specific for IgM (Southern Biotech), IgG1 (Southern Biotech) or IgE (Southern Biotech) overnight. Sera or supernatants were added and incubated for 2h at 37°C. After washing, alkaline phosphatase (AP)-conjugates of goat anti mouse IgM (Southern Biotech), IgG1 (Southern Biotech) and IgE (Southern Biotech) were incubated 1h at 37°C. After washing and addition of AP substrate (Sigma), absorbance was measured at 405 nm.

Anti-Ova specific antibodies produced *in vivo* after immunization were evaluated in sera from RHPS4 treated or untreated mice (10 mg/kg). Briefly, plates were coated with 10 µg/mL ovalbumin overnight. Sera were then incubated for 2h at 37°C and plates were treated as above.

ELISPOT

Specific IgG and IgM BCR anti-OVA antibody secreting cells were quantified by ELISPOT using splenocytes from mice sacrificed 9 days after immunization. Splenocytes were seeded in duplicate at a density starting at 5×10^5 /well, followed by 2-fold serial dilutions in culture medium in a 96-well MultiScreen HTS plate (Millipore) coated with 200 µg /well OVA. Cells were incubated overnight at 37 °C and then removed by washing with PBS/0.01 % Tween. Plates were incubated for 1 h at 37 °C with 1 µg/well of alkaline phosphatase-coupled anti-IgG or anti-IgM. After washing, 100 µL BCIP/NBT alkaline phosphatase substrate

(Millipore) was added. Plates were washed, dried, and images were taken with a NI-E microscope (Nikon) and analyzed for spot numbers using Nis-Ar software (Nikon).

Flow cytometry

Antibodies for staining were anti-mouse IgM-APC (Clone II/41, eBiosciences), CD19-APC-H7 (Clone 1D3, BD Biosciences), IgG1-BV421 (Clone A85-1, BD Biosciences), CD138-APC (Clone 281-2, BD Biosciences), B220-BV421 (Clone RA3-6B2, BD Biosciences). Apoptosis was evaluated using AnnexinV-FITC (BD Biosciences) according to the manufacturer's instructions and 20 μ L 7AAD (BD Pharmingen) was added 5 min prior to analysis.

Data were acquired on a Beckton Dickinson LSRII Fortessa cytometer and analyzed with the BD FACSDiva 6.1.3 and Infinicyt softwares.

Transcription analysis (RT-qPCR)

$I\mu$ -C μ , and $I\gamma$ 1-C γ 1 germline transcripts and $I\mu$ -C γ 1, $I\mu$ -C ϵ post switch transcripts were quantified. 4-day *in vitro* stimulated and RHPS4 treated B cells were collected and RNA was extracted for evaluation of post switch and germline transcripts. RNA was prepared using standard techniques. cDNA was synthesized using the High Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific). Quantitative PCR was performed using power SYBR green (Applied Biosystems) and specific oligonucleotides ($I\gamma$ 1 forward primer: 5' - GGCCCTTCCAGATCTTTGAG-3'; C γ 1 reverse primer: 5' - ATGGAGTTAGTTTGGGCAGCA-3'. $I\mu$ forward primer: 5' -ACCTGGGAATGTATGGTTGTGGCTT-3'. C μ reverse primer 5' -TCTGAACCTTCAAGGATGC TCTTG-3'. C ϵ reverse primer: 5-AGCGATGAATGGAGTAGC-3').

All transcripts were normalized to Glyceraldehyde-3-phosphate dehydrogenase (Gapdh) transcripts (reference Mm99999915-g1).

Binding of RHPS4 to S-region sequences.

Ability of RHPS4 (1 μ M or 2 μ M) to bind S-region derived DNA sequences was tested by fluorescence emission quenching. Various single-stranded oligonucleotides (1 μ M) mimicking G-rich portions of the S γ 1 region were compared to an oligothymidylate negative control (dT30). All measurements were performed in duplicate; fluorescence was collected in the 510-600 nm region on a TECAN Infinite M1000 Pro 96-well plate reader. Dye emission was normalized at each dye concentration by the intensity collected in the absence of DNA, partially (50-70%) quenched by binding to the G-rich S μ sequences, as previously described with other quadruplex-prone oligonucleotides^{E1}. All experiments were performed at room temperature in a lithium cacodylate buffer (pH 7) supplemented with 100 mM KCl to favor quadruplex formation. The DNA sequences are as follows (with G-blocks shown in color):

dT30	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
Smu2	TGGGGCAGGACTGCAGGGAAGTGGGGTATCAAGTAGAGGGA
Smu345	TGGGCTGAGCTGGGCAGGGCTGGGC
Smu381	TGGGCTGGGCTGAGCTGGGCTGGGCTGGGCTGGGCTGAGCGGGTCTGAGCGGGGC
Smu641	AGGCTGGGCTGGGCTGGGCTGAGCGGGGCTGAGCGGGGC
Smu1266	TGGGTTGGGCAGGGCTGGGCTGGGC
Smu1656	TGGGTTGGGCAAGGCTGGGCTGGGTTGAGTTGGGC

Amplification of S μ /S γ junctions and Ion torrent next generation sequencing NGS

DNA from LPS+IL-4 stimulated cells (treated with or without 1 μ M RHPS4) was extracted using classical phenol/chloroform protocol. S μ /S γ junctions were amplified in triplicate by

nested PCR with 100 ng DNA (Phusion HF polymerase; BioLabs) using the following primers: PCR1: Forward 5'- AGA GAC CTG CAG TTG AGG CC-3' and reverse 5'- TCA GGG AAR TAV CCY TTG ACC AGG CA-3', PCR 2: forward 5'-CCA GCC ACA GTA ATG ACC CAG-3' and reverse: 5'-CCA RKG GAT AGA CHG ATG GGG-3'. Each library was prepared using 200 ng PCR2 product. Barcoded libraries with 200-pb read lengths were prepared using Ion Xpress plus Fragment Library Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Each barcoded library was mixed in equal amounts and diluted to 100 pM. Libraries were run on an Ion PI v3 chip on the Ion Proton sequencer (Life Technologies). Data analysis was performed using CSReport^{E3}.

ChIP experiments

ChIP experiments were done on LPS+IL-4 stimulated CD43^{neg} spleen cells incubated with or without 1 μ M RHPS4. Briefly, 15×10^6 B-cells were cross-linked at room temperature for 15 min in 15 mL PBS with 1% formaldehyde. The reaction was quenched with 2.125M glycine. After lysis, chromatin was sonicated to 0.5–1 kb using a Vibracell 75043 (Thermo Fisher Scientific). Following 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 2h at 4 °C with 50 mL of 50% protein A/G slurry (0.2 mg / mL sheared salmon sperm DNA, 0.5 mg/ mL BSA, and 50% protein A/G; Sigma). Cell equivalents (1×10^6) were saved as input, and the remaining cell equivalents were incubated overnight with anti-AID rabbit polyclonal antibodies (kindly provided by Dr. P. Gearhart) or control antibodies. Immune complexes were precipitated by the addition of protein A/G. Cross-linking was reversed by overnight incubation (70 °C) in TE buffer with 0.02% SDS and chromatin was phenol/chloroform extracted. QPCR assays used for evaluating precipitated DNA from S μ , S γ 1 and S ϵ used the following primers:

- S μ Fw: 5'-TCTAAAATGCGCTAAACTGAGG-3'

- S_μRev: 5'-AGCGTAGCATAGCTGAGCTC -3'
- S_γ1Fw: 5'-GGGGGAGGAGATATCCAAGA -3'
- S_γ1Rev: 5'-CAGCTCTTTTGCAGGTCTTG-3'
- S_εFw: 5'-CTTGACCACCGAATGTCCTT -3'
- S_εRev: 5'-GATTCCTCTCCAGCCTCTCC-3'

Ovalbumin (OVA) induced allergic asthma

BALB/c female mice at 8 weeks of age were immunized i.p. on days 0, 7 and 14 with 20 μg ovalbumin grade V (Sigma France) emulsified in 2 mg aluminium hydroxide gel in a total volume of 200 μL. For control mice, 200 μL saline was injected.

Mice were challenged at days 21-24 with 10μg OVA by intra-tracheal administration to provoke allergic asthma with analysis of the allergic response at day 25. Control mice received saline.

Bronchoalveolar Lavage (BAL) and differential cell counts

BAL was performed by washing the lungs four times with 0.5 mL saline solution at room temperature. After centrifugation at 400 x g for 10 min at 4°C, the supernatant of the first lavage (cell-free BAL fluid) was stored at -80°C for cytokine analysis.

Cells were diluted with Turk's solution, counted and 200,000 cells were centrifuged onto microscopic slides (cytospin at 1,000 rpm for 10 min, at RT). Air-dried preparations were fixed and stained with Diff-Quik (#130832, Medion Diagnostics AG, Merz & Dade, Germany). Differential counts were made by oil immersion light microscopy. Two hundred cells were observed to determine the relative percentage of each cell type and absolute number of the differential cell count.

Lung homogenization and EPO activity

After bronchoalveolar lavage, the entire lung was perfused with isotonic solution through the right heart ventricle to flush the vascular content and lungs were frozen at -20°C until use.

Lungs were homogenized in 0.05M Tris-HCl, pH 8 with an Ultra Turrax T18 apparatus (Ultra Turrax, IKA), centrifuged and the supernatant was aliquoted. Following centrifugation, EPO activity was determined in the supernatant from lungs by colorimetric assay.

Following centrifugation, 100 μL of supernatants were placed in a plate with 50 μL of substrate solution, corresponding to 11 mL Tris HCl + 200mM O-phenylenediamine dihydrochloride pellets + 100 μL 30% H_2O_2). After 1 hour incubation at 37°C in a shaker, EPO activity was determined as absorbance at 490 nm against medium.

Lung histology

After BAL and lung perfusion, the left lobe was fixed in 4% buffered formaldehyde (#15225582, Fisher Scientific) for minimum 24h for standard microscopic analysis (Leica microscope). 3- μm sections were stained with Haematoxylin & Eosin (H&E) and Periodic Acid Schiff (PAS) using standard techniques.

Inflammatory cell including eosinophil infiltration was assessed by a semi-quantitative score on H&E sections (0-5), while mucus production was assessed on PAS stained sections with mucopolysaccharide at the apical pole of goblet cells of the bronchioles by comparable semi-quantitative score (0-5) by two observers independently.

Cytokine measurement

Th1/Th2 (IL-4, IL-5, IL-13 and IFN- γ) cytokines concentrations in BALF and lung homogenate were determined by Luminex immunoassay (Millipore) by using MagPix system (BioRad) according to manufacturer's instructions.

Sequence analysis:

Position of G4 DNA in the endogenous IgH locus was evaluated on the template and the non-template DNA strand from the C μ region to the 3' end of the locus, using the G4-hunter algorithm <http://bioinformatics.cruk.cam.ac.uk/G4Hunter/>.^{E2}

Sequence data from amplified CSR junctions were analyzed using CSReport^{E3}.

Statistical analysis:

Statistical tests were performed using GraphPad Prism (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001).

REFERENCES TO METHODS

- E1. Cheng, M.-K. et al. Antitumor polycyclic acridines. 20. Search for DNA quadruplex binding selectivity in a series of 8,13-dimethylquino[4,3,2-kl]acridinium salts: telomere-targeted agents. *J. Med. Chem.* 51, 963–975 (2008).
- E2. Bedrat, A., Lacroix, L. & Mergny, J.-L. Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.* 44, 1746–1759 (2016).
- E3. Boyer, F. et al. CSReport: A New Computational Tool Designed for Automatic Analysis of Class Switch Recombination Junctions Sequenced by High-Throughput Sequencing. *J. Immunol.* 198, 4148–4155 (2017).

SUPPLEMENTAL FIGURE LEGENDS

Figure E1. Occurrence of G4 DNA along mouse IgH constant locus.

Position of G4 DNA was evaluated on the non-template (positive scores) and the template (negative scores) DNA strand from the C_μ region to the 3' end of the locus, using the G4-hunter algorithm (<http://bioinformatics.cruc.cam.ac.uk/G4Hunter/>)^{E2}.

Figure E2. Effect of RHPS4 on cell growth and viability of mouse primary B cells.

(A) Mouse primary B cells stimulated *in vitro* by LPS-IL-4 for 4 days and treated with 0.1, 0.5, 1 or 2 μM RHPS4. Proliferation was measured by MTS assay. (B) Cell survival and death were analyzed by staining with AnnexinV-7AAD after 4 days *in vitro* culture as in A and in the presence of 1 μM RHPS4. Data represent mean % cells ± SEM for 6 mice. (C) Absolute numbers of plasmablasts in 4-day LPS-stimulated cultures, with or without RHPS4 treatment. Data represent mean % cell growth ± SEM for 5 mice. (D) Splenocyte viability from groups of mice injected *in vivo* daily with 10 mg/kg RHPS4 i.p. NS: not significant. **P*<0.05, ***P*<0.01, ****P*<0.001 compared with untreated cells.

Figure E3. Binding of RHPS4 to S-region sequences tested by fluorescence emission quenching.

RHPS4 was tested at 1 μM (hatched bars) and 2 μM (grey bars) concentrations, in the absence ("0") or presence of various oligonucleotides (1 μM) mimicking G-rich portions of the S_μ region or a single-stranded oligothymidylate control (dT30). All measurements were performed in duplicate. Fluorescence (collected in the 510-600 nm region) was collected on a TECAN Infinite M1000 Pro plate reader with 96-well plates. The emission of the dye was normalized at each dye concentration by the intensity collected in the absence of DNA partially (50-70%) quenched by binding to the G-rich S_μ sequences.

Figure E4. RHPS4 impeded CSR and affected AID-induced break localization.

CSR junctions from stimulated primary mouse B cells were amplified by PCR and sequenced by next generation sequencing. Counts of identified junctions (**A**), junction structures (**B**) and relative position of breaks in S γ 1 to AID hotspots (**C**) were obtained from CSReport sequence analysis. Data represent mean \pm SEM for 6 mice (A, C right panels) or one representative sample (B, C left panels). *P<0.05, **P<0.01 compared with untreated control cells.

Figure E5. Th cytokines and EPO activity in mice with OVA-induced allergic lung inflammation.

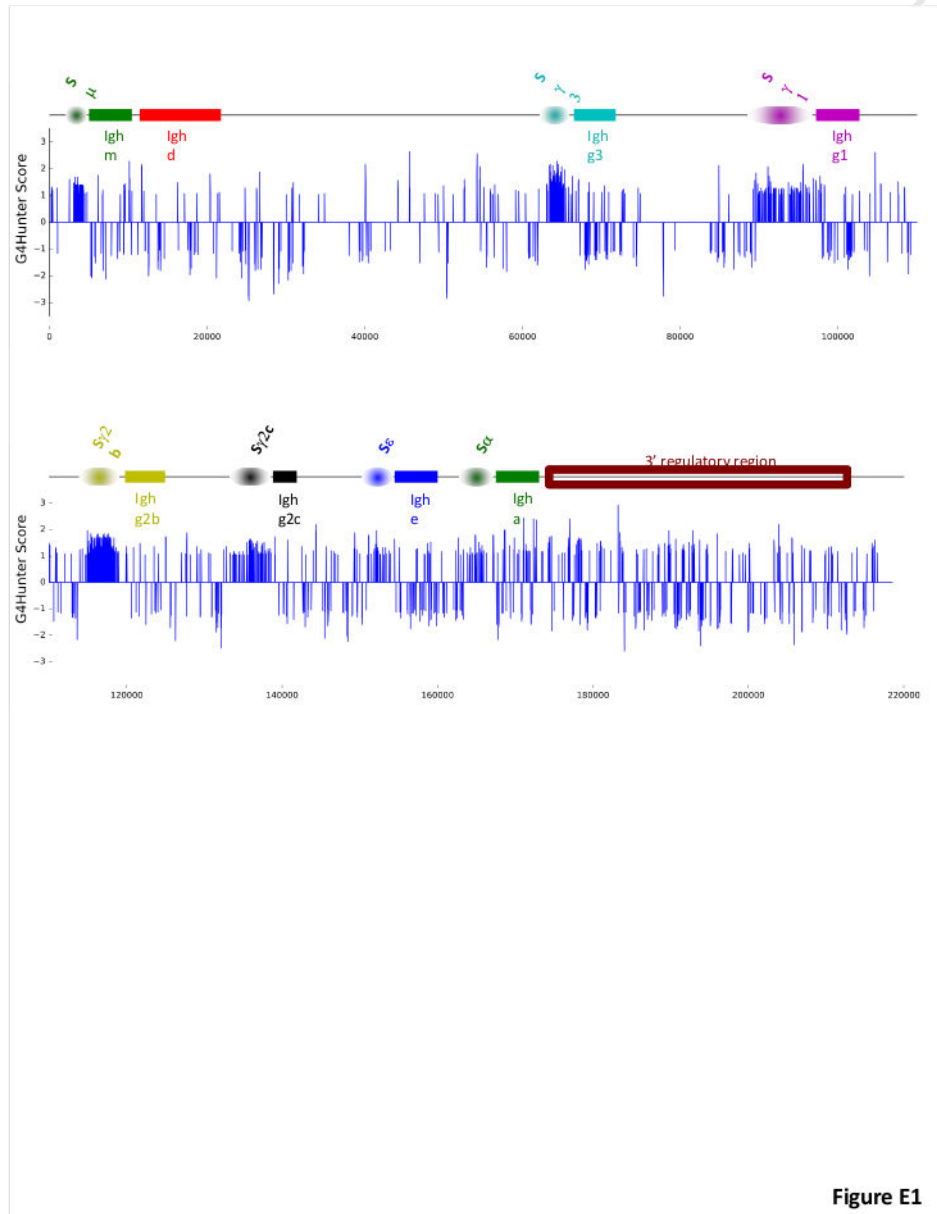
In an allergic asthma model after repeated *i.p.* and *i.n.* antigen administration, mice administered daily with steroids, and to a lesser extent mice receiving RHPS4, displayed (**A**) lower levels of IL-4, IL-5, IL-13 and IFN- γ in bronchoalveolar fluid (BALF) and (**B**) lower levels of IL-4, IL-5, IL-13 and IFN- γ in lung tissue.

Figure E6. Functional parameters of asthma and EPO activity in mice with OVA-induced allergic lung inflammation.

After repeated *i.p.* and *i.n.* antigen administration triggering allergic lung inflammation, mice receiving steroids or RHPS4 daily showed significantly improved (**A**) airway compliance (with both drugs) and (**B**) airway resistance (only significantly improved with steroids) by comparison to mice simply receiving antigen, and (**C**) had significantly lower lung eosinophil peroxidase activity (with both drugs).

Figure E7. Graphical abstract recapitulating the effects of RHPS4.

ACCEPTED MANUSCRIPT



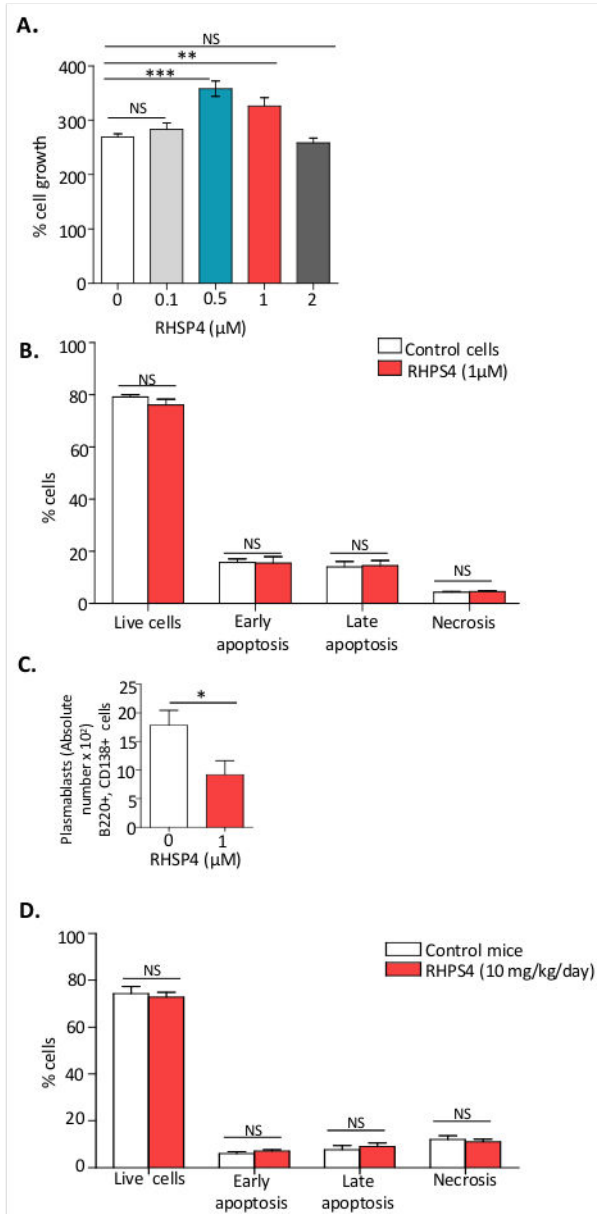
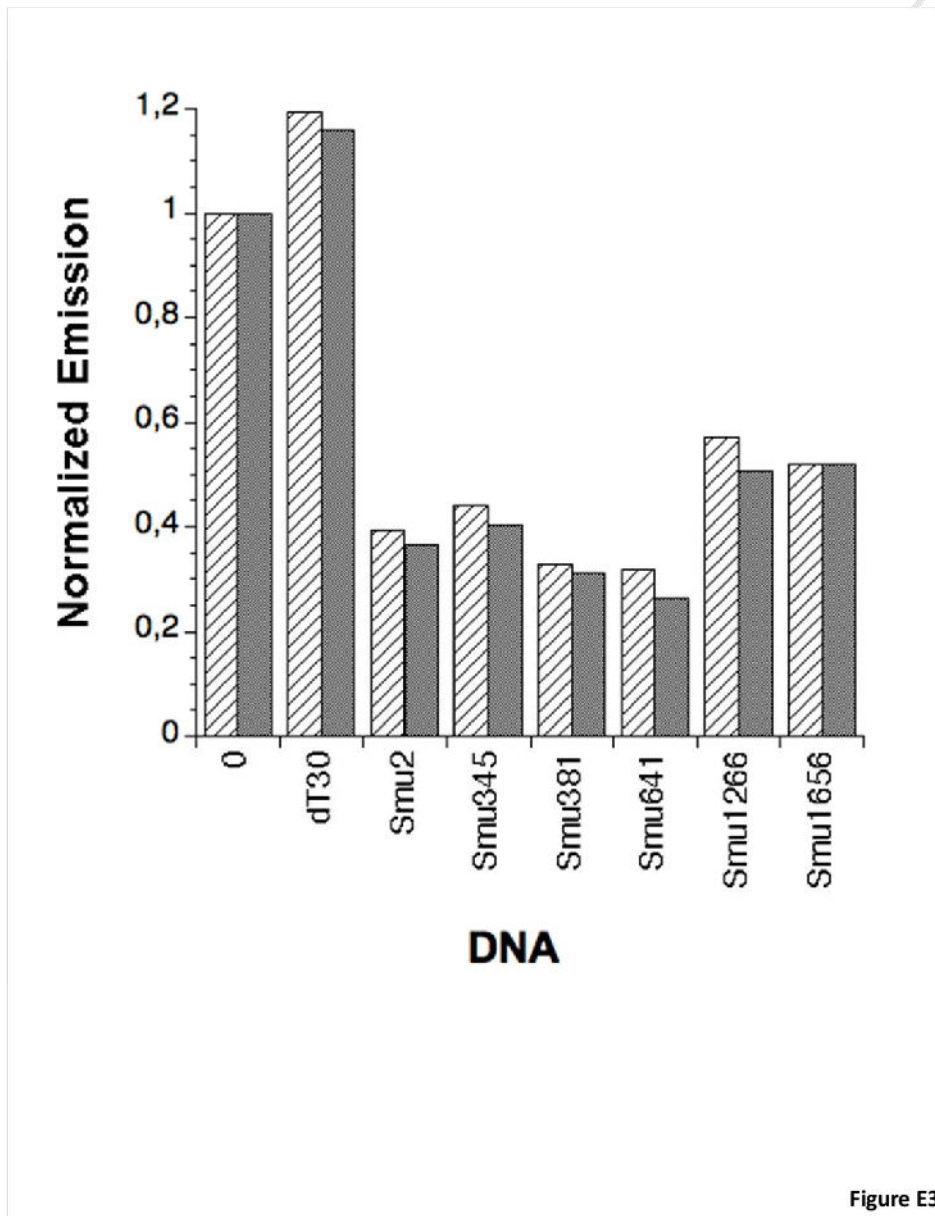
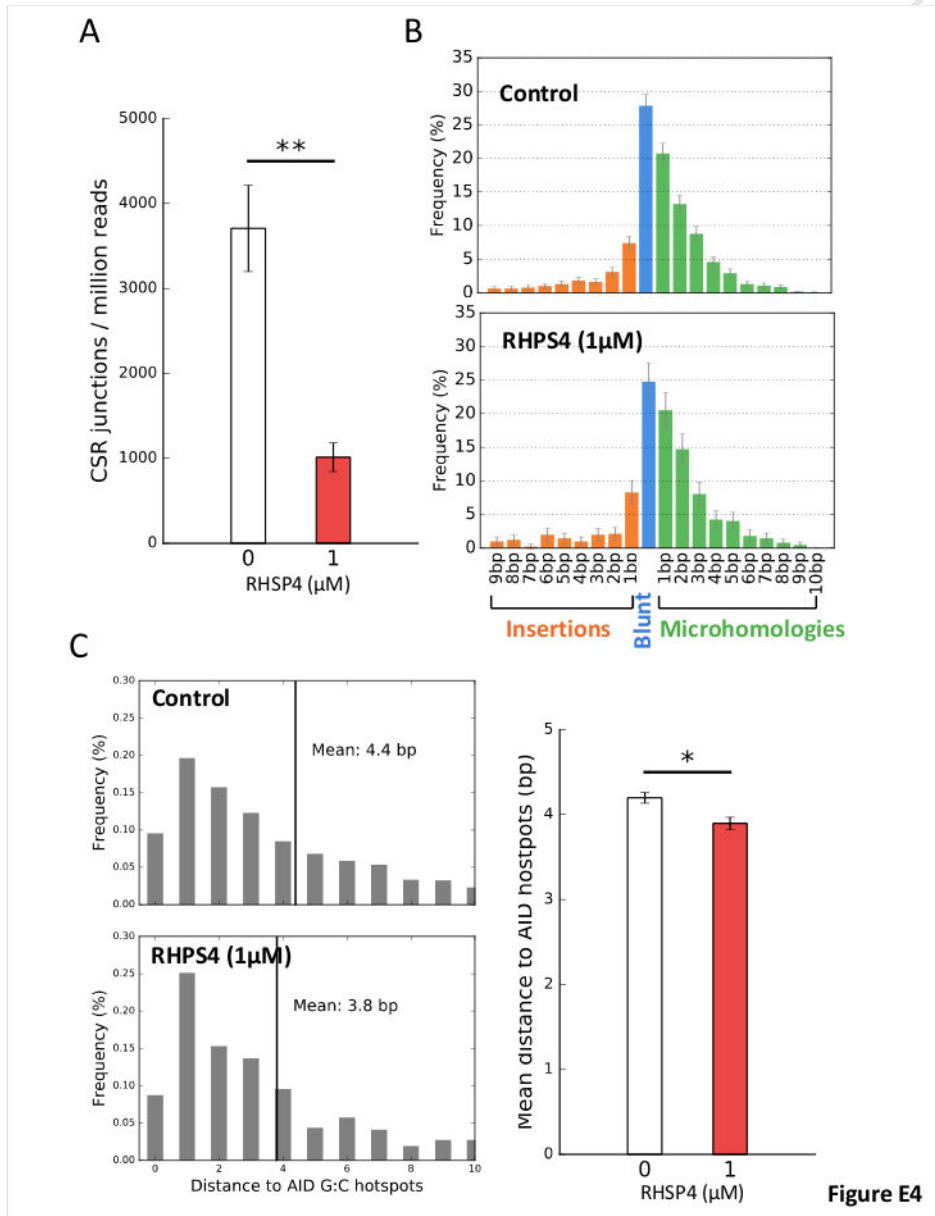
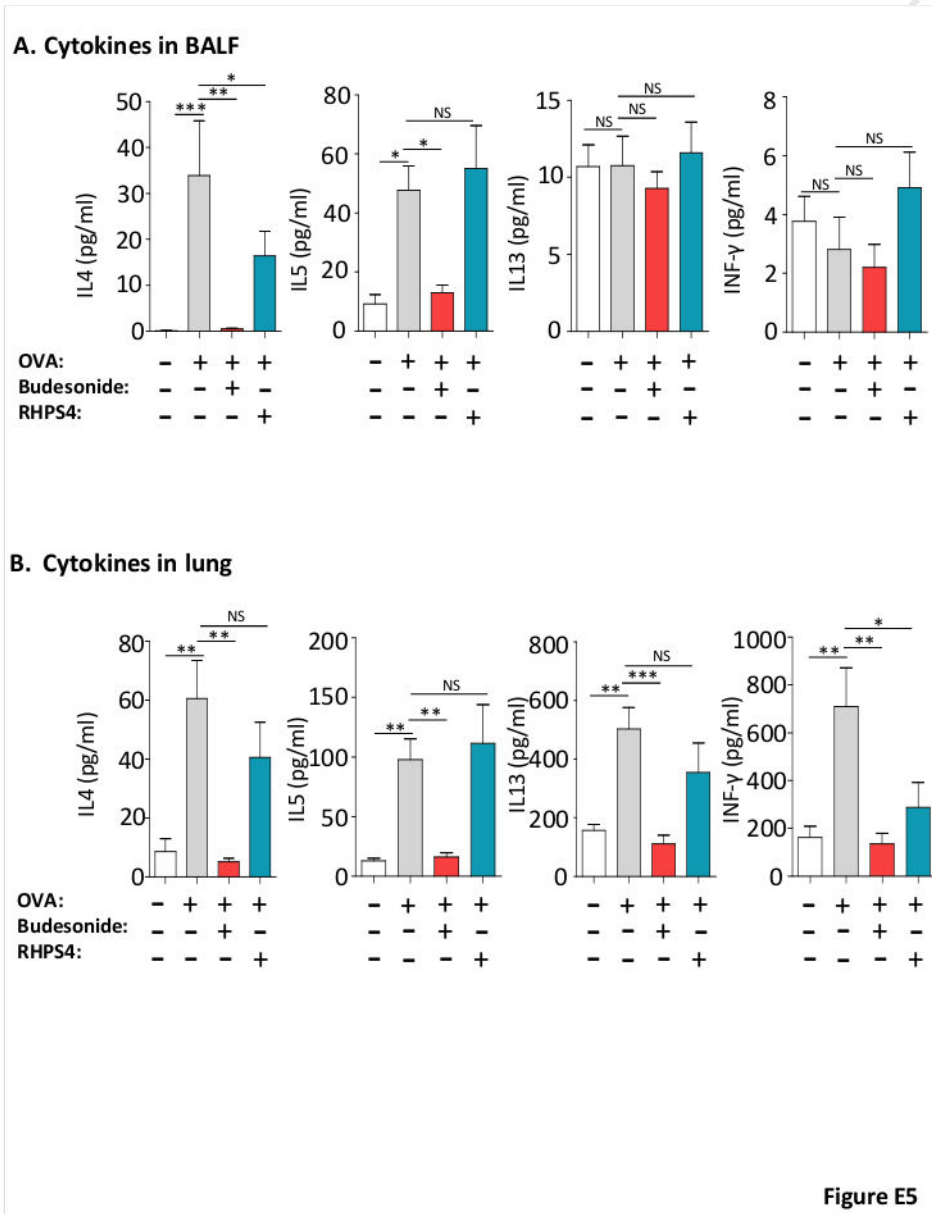


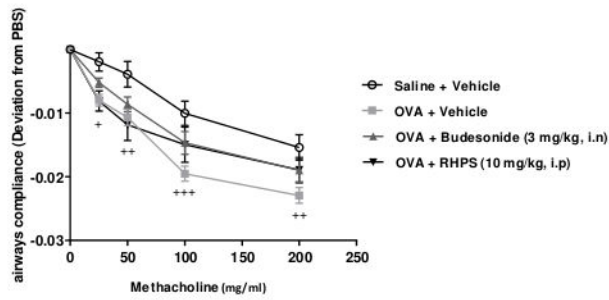
Figure E2



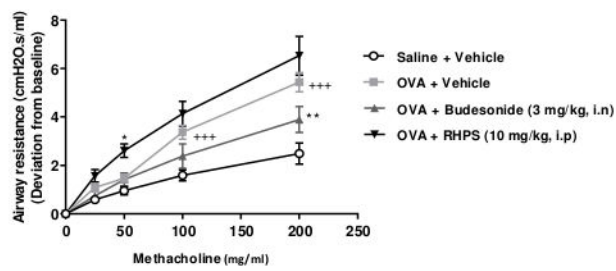




A. Airway compliance by plethysmography



B. Airway resistance by plethysmography



C. Lung eosinophil peroxidase activity

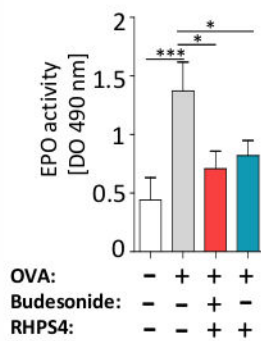
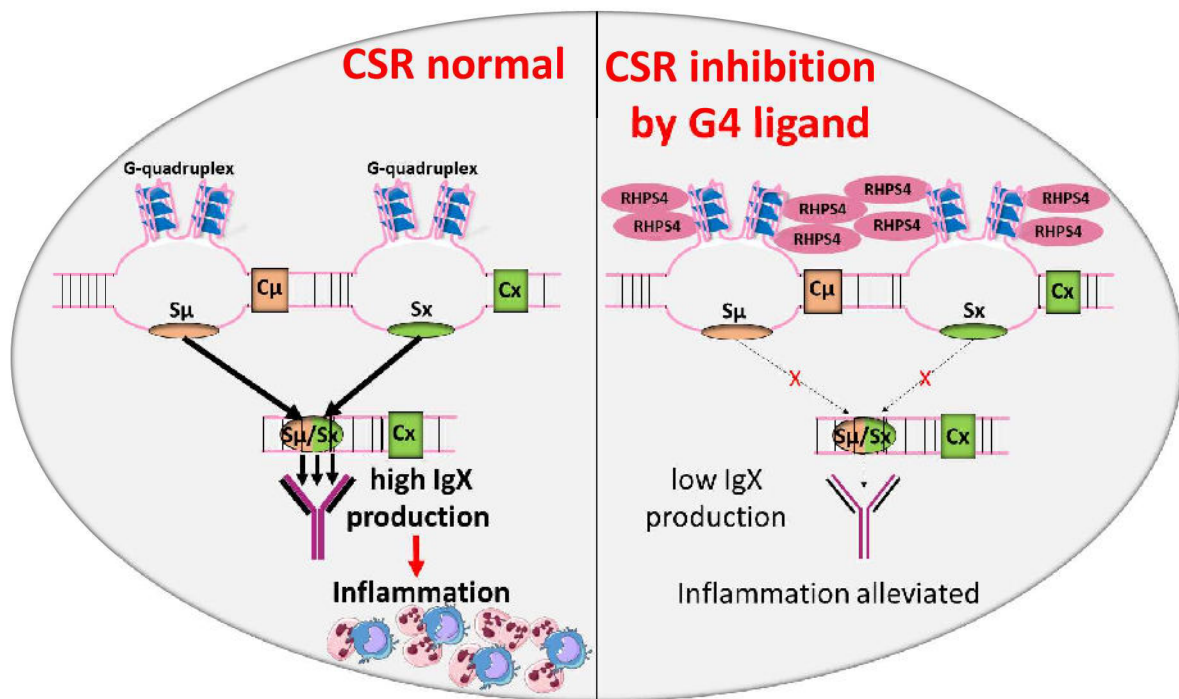


Figure E6



ACCEPTED M

Annexe 3 : (Revue)

Quand les gènes codant les anticorps s'adjoignent des fragments venus d'autres chromosomes.

When Immunoglobulin Genes Assemble with Bricks from Other Chromosomes.

Brice Laffleur, Zeinab Dalloul, Iman Dalloul, Sandrine Le Noir et Michel Cogné.

Commentaire scientifique accepté à Médecine/Sciences M/S en septembre 2016.

Cette revue décrit les remaniements précoces et tardives qui touchent le locus IgH durant le développement des cellules B. Le locus IgH, peut subir des recombinaisons en *Cis* avec le même chromosome mais aussi en *Trans* avec le chromosome homologues ou non homologue. Ces recombinaisons en *Trans* entre des chromosomes non homologues ont été identifiées chez des patients atteints de paludisme, dont le gène LAIR1 (*Leukocyte Associated Immunoglobulin like Receptor 1*) situé sur le chromosome 19 est trouvé introduit entre le segment V et DJ du locus IgH situé sur le chromosome 14.

Dans cette revue, j'ai participé à dessiner les figures.



RÉFÉRENCES

- Thion MS, McGuire JR, Sousa CM, et al. Unraveling the role of huntingtin in breast cancer metastasis. *J Natl Cancer Inst* 2015 ; 107 : djv208.
- Sorensen SA, Fenger K, Olsen JH. Significantly lower incidence of cancer among patients with Huntington disease: An apoptotic effect of an expanded polyglutamine tract? *Cancer* 1999 ; 86 : 1342-6.
- Ji J, Sundquist K, Sundquist J. Cancer incidence in patients with polyglutamine diseases: a population-based study in Sweden. *Lancet Oncol* 2012 ; 13 : 642-8.
- Charafe-Jauffret E, Chaffanet M, Bertucci F, et al. Les cancers du sein : vers un modèle cellulaire et moléculaire intégré. *Med Sci (Paris)* 2007 ; 23 : 626-32.
- Tokes AM, Szasz AM, Juhasz E, et al. Expression of tight junction molecules in breast carcinomas analysed by array PCR and immunohistochemistry. *Pathol Oncol Res* 2012 ; 18 : 593-606.
- Hoover KB, Liao SY, Bryant PJ. Loss of the tight junction MAGUK ZO-1 in breast cancer: relationship to glandular differentiation and loss of heterozygosity. *Am J Pathol* 1998 ; 153 : 1767-73.
- Martin TA, Watkins G, Mansel RE, et al. Loss of tight junction plaque molecules in breast cancer tissues is associated with a poor prognosis in patients with breast cancer. *Eur J Cancer* 2004 ; 40 : 2717-25.
- Lo Sardo V, Zuccato C, Gaudenzi G, et al. An evolutionary recent neuroepithelial cell adhesion function of huntingtin implicates ADAM10-Ncadherin. *Nat Neurosci* 2012 ; 15 : 713-21.

NOUVELLE

Quand les gènes codant les anticorps s'adjoignent des fragments venus d'autres chromosomes

Brice Laffleur¹, Zeinab Dalloul², Iman Dalloul², Sandrine Le Noir², Michel Cogne^{2,3}

¹Department of immunology, Columbia university, New York, États-Unis ;

²Université de Limoges, CNRS UMR 7276, Contrôle de la réponse immune et lymphoprolifération, faculté de médecine, 2, rue du Docteur Marcland, 87025 Limoges, France ;

³Institut universitaire de France, Paris, France.

cogne@unilim.fr

Diversité du système immunitaire et recombinaisons chromosomiques

L'immunité met en œuvre une multiplicité de cellules et molécules aptes à reconnaître un immense répertoire de structures moléculaires portées par les antigènes. Ceci permet à la fois la « tolérance du soi » en condition basale et une réponse à diverses agressions grâce à une immunité contre le « non-soi » (comme des microorganismes) ou le « soi modifié » (induit par des mutations, un cancer, etc.). Chez les invertébrés, l'immunité innée repose sur un répertoire limité de récepteurs invariants reconnaissant un spectre restreint d'antigènes. Elle a l'avantage de préexister à l'arrivée de tout antigène étranger mais elle présente peu de spécificité et d'affinité, et aucune adaptabilité. Chez les mammifères, s'ajoute une immunité adaptative qui est, elle, capable de lier une immense diversité d'antigènes avec spécificité et affinité via deux grands répertoires très diversifiés de cellules (les lymphocytes T et B) et de récepteurs ou molécules solubles : les récepteurs TCR (*T cell receptor*) et BCR (*B cell receptor*) ou les immunoglobulines (Ig) solubles. L'économie des gènes

codant ces récepteurs permet de générer leur diversité de façon combinatoire, en réassociant des fragments de gènes à la manière d'un patchwork pour créer dans une cellule la combinaison codant une Ig ou un TCR donné. Des réarrangements chromosomiques génèrent ainsi une variété quasi-infinie de récepteurs à partir de seulement quelques dizaines de gènes ou briques élémentaires (*Figure 1*). Alors que ces mécanismes semblaient connus en détails, un type de réarrangement très inattendu vient d'être décrit comme permettant la génération d'anticorps de large réactivité contre le parasite responsable du paludisme, *Plasmodium falciparum* [1], remettant certains dogmes en questions...

Génération de la diversité des récepteurs à l'antigène

Les récepteurs à l'antigène, TCR et BCR, des lymphocytes T et B sont codés via la mécanique moléculaire commune de la « recombinaison V(D)J ». Celle-ci survient avant la rencontre avec l'antigène, dans le thymus pour les lymphocytes T, et dans la moelle osseuse pour les B. Elle assemble des segments de gènes

nommés V (variable), D (diversité) et J (jonction) (*Figure 1A*). Le brassage combinatoire de ces quelques dizaines de gènes s'opère grâce aux enzymes spécialisées RAG (*recombination activating gene*) qui coupent l'ADN aux extrémités des segments V(D)J, au niveau de signaux dits RSS (*recombination signal sequences*). Les mécanismes classiques de réparation de l'ADN permettent ensuite de ressouder deux segments entre eux. À la diversité combinatoire, ces mécanismes de réparation ajoutent une certaine « imprécision » (par l'ajout ou le retrait de quelques bases au point de cassure qui a été réparé), augmentant ainsi la diversité des séquences générées (diversité dite « jonctionnelle ») [2, 3]. L'assemblage unique par chaque clone T ou B de tels segments « V(D)J », génère ainsi des populations lymphocytaires polyclonales qui expriment globalement un répertoire de récepteurs dont la diversité est quasi-infinie. Parmi ces millions de clones (chacun arborant un récepteur de spécificité définie), certains seront activés par la rencontre d'un antigène spécifique et participeront donc à la réponse immune.

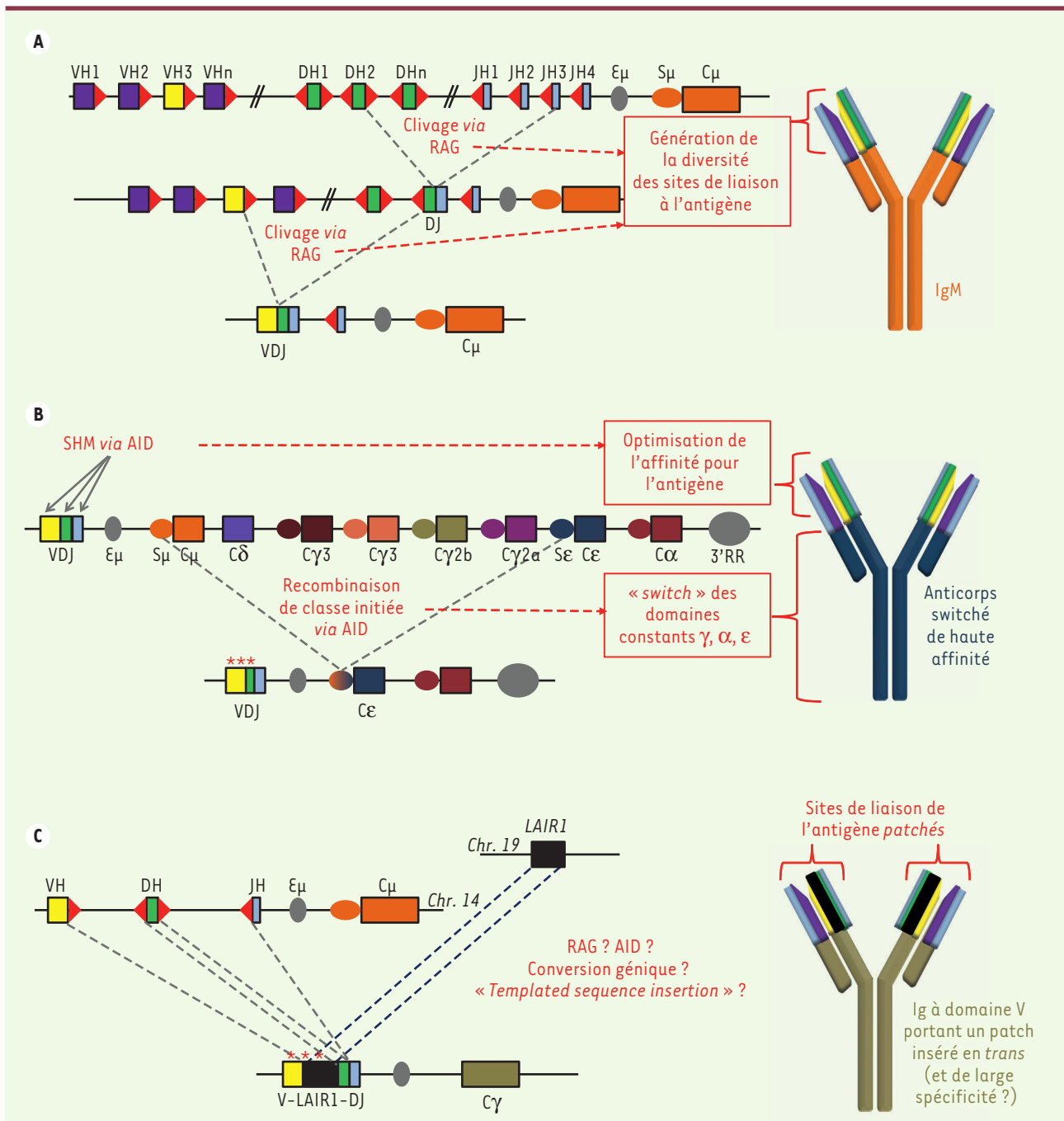


Figure 1. Mécanique des remaniements des gènes d'immunoglobulines. **A.** Génération initiale de la diversité des immunoglobulines (Ig) par RAG (recombination activating gene) et recombinaison VDJ au sein du locus des chaînes lourdes d'Ig (IgH). Les segments V, D, J et constant du locus IgH sont respectivement représentés en violet et jaune, vert, bleu et orange ; les RSS (recombination signal sequences) sont schématisés par des triangles rouges. La partie droite représente une IgM après recombinaison VDJ. **B.** Optimisation des Ig sous l'effet d'AID (activation induced cytidine deaminase) après rencontre de l'antigène, par hypermutation somatique et recombinaison de commutation de classe (switch ou CSR [class switch recombination]) du locus IgH. Les régions switch et les gènes constants sont respectivement représentés par des ronds et des carrés. Les étoiles correspondent aux mutations induites par AID sur le segment VDJ. La partie basse du schéma représente le locus IgH après CSR ; la partie droite représente une Ig après SHM (somatic hypermutation) et CSR, dans l'exemple du switch IgE. **C.** Insertion par recombinaison en « trans » d'une séquence non-Ig LAIR1 au sein d'un gène réarrangé VDJ du locus IgH et génération d'un domaine variable porteur d'un patch. Schéma représentatif d'une insertion du gène LAIR1 (leukocyte associated immunoglobulin like receptor 1) (représenté en noir) lors d'une recombinaison (VDJ-like ?) du locus IgH. Le schéma de droite représente une Ig avec un site de fixation à l'antigène comprenant l'addition d'un patch LAIR1.



Les cellules T orchestrent ces réponses et peuvent détruire des cellules cibles ; les cellules B pourront sécréter leurs BCR sous la forme d'Ig solubles, les anticorps. Une mémoire immunitaire sera également générée permettant une réponse accélérée en cas de nouvelle rencontre avec le même antigène (vaccination).

Maturation d'affinité et commutation de classe des immunoglobulines

Afin d'adapter la réponse immune après la rencontre avec l'antigène, la lignée B est capable de prolonger le remaniement des gènes d'Ig grâce à l'enzyme AID (*activation induced cytidine deaminase*). Celle-ci initie deux processus : l'hypermutation somatique (SHM, pour *somatic hypermutation*) et la recombinaison de classe (CSR, pour *class switch recombination*) (Figure 1B) (pour revue voir [4]). Un segment VDJ préalablement assemblé pourra alors être modifié et son site de liaison à l'antigène amélioré par SHM. Un processus de sélection cellulaire favorise les cellules présentant le BCR ayant la plus haute affinité vis-à-vis de l'antigène. Plusieurs « classes » de BCR et d'Ig (IgM, puis IgG, IgA ou IgE, avec différents domaines constants) sont en outre successivement produites par un clone B (Figure 1B). Cette « commutation isotypique » ou « switch » résulte de la CSR entre des séquences d'ADN (régions « switch » des gènes d'Ig) dont le ciblage par AID crée des cassures double brin. Un dernier type AID-dépendant de recombinaison est la LSR (*locus suicide recombination*), une variante du switch qui supprime la totalité des gènes constants IgH (à l'origine de la chaîne lourde des Ig) et entraîne donc la perte du BCR et l'apoptose [5].

Interactions cis- et trans-chromosomiques

Physiologiquement, dans l'immense majorité des cas, ces cassures programmées du génome réassocient *in fine* des fragments de gènes qui sont des partenaires légitimes (Ig-Ig dans les B, TCR-TCR dans les T). Les *trans*-recom-

binaisons avec d'autres chromosomes sont considérées comme rarissimes, relevant du champ de la pathologie. Les remaniements lymphoïde-spécifiques physiologiques du génome nécessitent donc un contrôle strict et leurs erreurs peuvent s'avérer dramatiques (comme la translocation d'un oncogène, à l'origine de leucémies ou de lymphomes) [10] (→).

L'expression et la localisation nucléaire des enzymes RAG et AID ainsi que leur accès à la chromatine sont ainsi régulées dans le temps et l'espace, strictement dans les lignées lymphoïdes, et à des stades spécifiques de leur développement. Les chromosomes impliqués sont alors relocalisés dans des compartiments nucléaires particuliers qu'on peut dénommer « *recombination factories* »¹. Différents activateurs transcriptionnels des gènes d'Ig ou de TCR orchestrent leurs remaniements en rapprochant les séquences cibles de recombinaison, créant d'éventuelles boucles chromosomiques lors de recombinaison « en cis » (sur un même chromosome). Les recombinaisons V(D)J sont réputées survenir quasi-exclusivement en cis lorsqu'elles surviennent entre des partenaires « légitimes » (c'est-à-dire entre segments V, D et J authentiques). Au contraire, les réarrangements de switch (qui surviennent de façon simultanée sur les deux allèles) impliquent fréquemment des recombinaisons inter-alléliques « en trans » (sortes de translocations légitimes entre chromosomes de la même paire) [6].

Un nouveau mécanisme de génération d'anticorps en trans

Le nouveau type de remaniement décrit par Tan *et al.* s'inscrit dans un contexte d'interactions *trans*-chromosomiques, mais avec la particularité d'impliquer, au moins partiellement, les recombinaisons V(D)J et d'apparier les locus d'Ig avec un gène autre situé sur un chro-

mosome non homologue [1]. De façon étonnante, elles semblent cependant physiologiques puisqu'elles aboutissent à la production d'une molécule qui porte la plupart des caractéristiques d'un anticorps fonctionnel présentant une réactivité large et qui aurait vraisemblablement un rôle actif dans la réponse immune anti-plasmodiale.

Ce nouveau type d'anticorps a été découvert chez deux individus différents, à l'occasion d'une étude portant sur les anticorps de réactivité large qui sont présents chez des sujets impaludés (le paludisme se traduit notamment par la présence de *P. falciparum* infectant les érythrocytes et induisant alors l'expression d'antigènes du parasite à la surface des hématies infectées). Alors que les malades peuvent présenter des accès sévères de paludisme en début d'infection ou de réinfection, les personnes vivant en zone impaludée développent une immunité et des anticorps dirigés contre l'isolat de parasite à l'origine de leur infection. Quelques sujets développent des anticorps dits « de large réactivité » qui sont capables d'agglutiner des hématies infectées par de multiples isolats indépendants de *P. falciparum*. L'étude de leur immunité présente donc un intérêt particulier, potentiellement révélateur de stratégies naturelles d'immunité anti-plasmodiale large. Ainsi, après avoir criblé des centaines d'échantillons, Tan *et al.* ont identifié quelques donneurs sains dont les anticorps sont capables de neutraliser des érythrocytes infectés par 6 isolats différents du parasite et d'entraîner leur élimination par phagocytose Ig-dépendante [1]. Les lymphocytes B de ces donneurs ont été immortalisés afin de déterminer la séquence de ces anticorps protecteurs de large réactivité. De façon totalement inattendue, le séquençage a révélé que le domaine de liaison à l'antigène de certains de ces anticorps était codé (chez deux patients indépendants) par un gène composite qui incluait une large portion étrangère aux

(→) Voir la Synthèse de S. Gisselbrecht, *m/s n° 2, février 2003, page 201*

¹ Usines à recombinaison.

gènes normaux d'Ig (qui sont situés sur le chromosome 14 humain) identifiable comme provenant du chromosome 19 mais recombinée et insérée entre les segments V et DJ. Ce segment de chromosome 19 correspond à une partie du gène *LAIR1* (*leukocyte associated immunoglobulin like receptor 1*), un récepteur inhibiteur (à domaines ITIM [*immunoreceptor tyrosine-based inhibition motif*]) appartenant à la superfamille des Ig exprimé par les cellules B et liant, normalement, le collagène. Une séquence codant 98 acides aminés de *LAIR1* est ainsi exprimée comme une sorte de « patch » inséré au sein d'un domaine variable d'Ig et permet, après hypermutation somatique, d'y lier non pas le collagène, mais des épitopes de la famille multigénique *RIFIN* (la plus large famille d'antigènes variants de *Plasmodium*, qui sont exprimés à la surface des érythrocytes infectés et impliqués en tant qu'adhésines dans certaines formes sévères du paludisme). De plus, une fois cette séquence insérée au sein du gène VDJ, elle subit comme un VDJ classique l'hypermutation permettant d'augmenter son affinité pour l'antigène plasmodial *RIFIN* (et de supprimer l'affinité de *LAIR1* pour le collagène) (*Figure 1C*).

Conclusion

Ce nouveau mécanisme démontre la capacité physiologique de nos cellules B à introduire au sein d'un gène d'Ig une séquence étrangère (non-Ig) afin de créer un néo-site de liaison à l'antigène. Le séquençage haut-débit devrait permettre d'évaluer la fréquence de telles insertions d'un patch *LAIR1* dans des populations saines et éventuellement exposées à des infections chroniques. D'autres patchs de séquences non-Ig ainsi insérés pourraient également être découverts et il reste donc à savoir si le cas de *LAIR1*, pour les anticorps anti-*Plasmodium*, restera ou non un exemple unique d'anticorps fonctionnel naturel incluant un patch « non-Ig ».

Le mécanisme de ce remaniement génique reste à préciser, mais il porte au moins la signature de RAG puisque l'insertion se fait au niveau du point de jonction entre les segments V et DJ. De plus, la séquence de *LAIR1* se révèle flanquée d'introns incluant des séquences qui évoquent des RSS cryptiques. Une *trans*-recombinaison simple médiée par RAG semble exclue car les cellules B des clones concernés ont conservé 2 chromosomes 19 portant des copies intactes du gène *LAIR1*. Ceci évoquerait donc plutôt un phénomène de type conversion génique, ou encore l'insertion au point de cassure d'un gène d'Ig d'un pré-ARN messager *LAIR1* rétrotranscrit (ce type de réparation mutagène des cassures double brin avec insertion d'une séquence rétrotransposée a récemment été décrit et dénommé « *templated sequence insertion* ») [7]. Pour les anticorps anti-*Plasmodium*, la complexité du phénomène décrit est encore accrue par l'implication claire d'AID dans les mutations secondaires du gène *V-LAIR1-DJ* réarrangé et du patch *LAIR1*. L'infection palustre entraîne une stimulation B chronique et une expression accrue d'AID, pouvant favoriser des translocations [8]. Il est possible qu'elle suscite également la *trans*-recombinaison de *LAIR1* au sein d'un segment VDJ, puis son hypermutation. Dans ce contexte, on se rappellera aussi que même s'ils échappent au dogme et au schéma classique de la maturation B, des épisodes de co-expression de RAG et AID ont été démontrés [9].

Alors que ce processus est sans doute plus l'exception que la règle, il dévoile cependant des aptitudes insoupçonnées des lymphocytes B à poursuivre encore plus loin l'extrême diversification de leur répertoire afin d'optimiser et d'adapter les réponses immunes aux antigènes rencontrés, par l'usage d'un arsenal dont la description vient de s'enrichir d'une nouvelle arme ! Il est frappant de noter que cette description survient au cours de la caractérisa-

tion d'anticorps d'un grand intérêt potentiel en immunothérapie, actuellement traqués dans le cadre de la protection contre divers agents infectieux (VIH [virus de l'immunodéficience humaine], *Norovirus*, *Influenza*, *P. falciparum*, etc.), les anticorps neutralisants à vaste spectre de réactivité [11] (→). ♦

(→) Voir la Nouvelle de T. Bruel et al. sur les anticorps bNAb, page 671 de ce numéro

When immunoglobulin genes assemble with bricks from other chromosomes

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Tan J, Pieper K, Piccoli L, et al. A *LAIR1* insertion generates broadly reactive antibodies against malaria variant antigens. *Nature* 2016 ; 529 : 105-9.
2. Nishana M, Raghavan SC. Role of recombination activating genes in the generation of antigen receptor diversity and beyond. *Immunology* 2012 ; 137 : 271-81.
3. Lieber MR, Yu K, Raghavan SC. Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. *DNA Repair* 2006 ; 5 : 1234-45.
4. Laffleur B, Denis-Lagache N, Péron S, et al. AID-induced remodeling of immunoglobulin genes and B cell fate. *Oncotarget* 2014 ; 5 : 1118-31.
5. Péron S, Laffleur B, Denis-Lagache N, et al. AID-driven deletion causes immunoglobulin heavy chain locus suicide recombination in B cells. *Science* 2012 ; 336 : 931-4.
6. Laffleur B, Bardet SM, Garot A, et al. Immunoglobulin genes undergo legitimate repair in human B cells not only after cis- but also frequent trans-class switch recombination. *Genes Immun* 2014 ; 15 : 341-6.
7. Onozawa M, Zhang Z, Kim YJ, et al. Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. *Proc Natl Acad Sci USA* 2014 ; 111 : 7729-34.
8. Robbiani DF, Deroubaix S, Feldhahn N, et al. *Plasmodium* infection promotes genomic instability and AID-dependent B cell lymphoma. *Cell* 2015 ; 162 : 727-37.
9. Wang JH, Gostissa M, Yan CT, et al. Mechanisms promoting translocations in editing and switching peripheral B cells. *Nature* 2009 ; 460 : 231-6.
10. Gisselbrecht S. Oncogènes et leucémies : historique et perspectives. *Med Sci (Paris)* 2003 ; 19 : 201-10.
11. Bruel T, Mouquet H, Schwartz O. Des anticorps qui détruisent les cellules infectées par le VIH-1. *Med Sci (Paris)* 2016 ; 32 : 671-4.

Références bibliographiques

- Abdelrasoul, H., Werner, M., Setz, C.S., Okkenhaug, K., and Jumaa, H. (2018). PI3K induces B-cell development and regulates B cell identity. *Scientific Reports* 8, 1327.
- Agrawal, S., Smith, S. a. B.C., Tangye, S.G., and Sewell, W.A. (2013). Transitional B cell subsets in human bone marrow. *Clin. Exp. Immunol.* 174, 53–59.
- Akoulitchev, S., Chuikov, S., and Reinberg, D. (2000). TFIID is negatively regulated by cdk8-containing mediator complexes. *Nature* 407, 102–106.
- Allen, B.L., and Taatjes, D.J. (2015). The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* 16, 155–166.
- Allman, D., Lindsley, R.C., DeMuth, W., Rudd, K., Shinton, S.A., and Hardy, R.R. (2001). Resolution of three nonproliferative immature splenic B cell subsets reveals multiple selection points during peripheral B cell maturation. *J. Immunol.* 167, 6834–6840.
- Allman, D.M., Ferguson, S.E., Lentz, V.M., and Cancro, M.P. (1993). Peripheral B cell maturation. II. Heat-stable antigen(hi) splenic B cells are an immature developmental intermediate in the production of long-lived marrow-derived B cells. *J. Immunol.* 151, 4431–4444.
- Aupetit, C., Drouet, M., Pinaud, E., Denizot, Y., Aldigier, J.C., Bridoux, F., and Cogné, M. (2000). Alleles of the alpha immunoglobulin gene 3' enhancer control evolution of IgA nephropathy toward renal failure. *Kidney Int.* 58, 966–971.
- Bahjat, M., and Guikema, J.E.J. (2017). The Complex Interplay between DNA Injury and Repair in Enzymatically Induced Mutagenesis and DNA Damage in B Lymphocytes. *Int J Mol Sci* 18.
- Banerji, J., Olson, L., and Schaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* 33, 729–740.
- Bébin, A.-G., Carrion, C., Marquet, M., Cogné, N., Lecardeur, S., Cogné, M., and Pinaud, E. (2010). In vivo redundant function of the 3' IgH regulatory element HS3b in the mouse. *J. Immunol.* 184, 3710–3717.
- Bergman, Y., and Cedar, H. (2004). A stepwise epigenetic process controls immunoglobulin allelic exclusion. *Nat. Rev. Immunol.* 4, 753–761.
- Bhattacharya, P., Wuerffel, R., and Kenter, A.L. (2010). Switch region identity plays an important role in Ig class switch recombination. *J. Immunol.* 184, 6242–6248.
- Birshtein, B.K. (2014). Epigenetic Regulation of Individual Modules of the immunoglobulin heavy chain locus 3' Regulatory Region. *Front Immunol* 5, 163.
- Bojarczuk, K., Bobrowicz, M., Dwojak, M., Miazek, N., Zapala, P., Bunes, A., Siernicka, M., Rozanska, M., and Winiarska, M. (2015). B-cell receptor signaling in the pathogenesis of lymphoid malignancies. *Blood Cells Mol. Dis.* 55, 255–265.
- Borzillo, G.V., Cooper, M.D., Kubagawa, H., Landay, A., and Burrows, P.D. (1987). Isotype switching in human B lymphocyte malignancies occurs by DNA deletion: evidence for nonspecific switch recombination. *J. Immunol.* 139, 1326–1335.

- Bosma, G.C., Kim, J., Urich, T., Fath, D.M., Cotticelli, M.G., Ruetsch, N.R., Radic, M.Z., and Bosma, M.J. (2002). DNA-dependent protein kinase activity is not required for immunoglobulin class switching. *J. Exp. Med.* *196*, 1483–1495.
- Braikia, F.-Z., Conte, C., Moutahir, M., Denizot, Y., Cogné, M., and Khamlichi, A.A. (2015). Developmental Switch in the Transcriptional Activity of a Long-Range Regulatory Element. *Mol. Cell. Biol.* *35*, 3370–3380.
- van der Burg, M., Tümkaya, T., Boerma, M., de Bruin-Versteeg, S., Langerak, A.W., and van Dongen, J.J. (2001). Ordered recombination of immunoglobulin light chain genes occurs at the IGK locus but seems less strict at the IGL locus. *Blood* *97*, 1001–1008.
- Carlsten, J.O.P., Zhu, X., and Gustafsson, C.M. (2013). The multitasking Mediator complex. *Trends Biochem. Sci.* *38*, 531–537.
- Chaturvedi, A., Martz, R., Dorward, D., Waisberg, M., and Pierce, S.K. (2011). Endocytosed BCRs sequentially regulate MAPK and Akt signaling pathways from intracellular compartments. *Nat. Immunol.* *12*, 1119–1126.
- Chaudhuri, J., and Alt, F.W. (2004). Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat. Rev. Immunol.* *4*, 541–552.
- Chauveau, C., and Cogné, M. (1996). Palindromic structure of the IgH 3' locus control region. *Nat. Genet.* *14*, 15–16.
- Cianci, R., Giambra, V., Mattioli, C., Esposito, M., Cammarota, G., Scibilia, G., Magazzù, G., Orlando, A., Sandri, G., Bianchi, L., et al. (2008). Increased frequency of Ig heavy-chain HS1,2-A enhancer *2 allele in dermatitis herpetiformis, plaque psoriasis, and psoriatic arthritis. *J. Invest. Dermatol.* *128*, 1920–1924.
- Clark, M.R., Mandal, M., Ochiai, K., and Singh, H. (2014). Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nat. Rev. Immunol.* *14*, 69–80.
- Cogné, M., Lansford, R., Bottaro, A., Zhang, J., Gorman, J., Young, F., Cheng, H.L., and Alt, F.W. (1994). A class switch control region at the 3' end of the immunoglobulin heavy chain locus. *Cell* *77*, 737–747.
- Collins, A.M., and Jackson, K.J.L. (2018). On being the right size: antibody repertoire formation in the mouse and human. *Immunogenetics* *70*, 143–158.
- Conaway, R.C., and Conaway, J.W. (2011). Origins and activity of the Mediator complex. *Semin. Cell Dev. Biol.* *22*, 729–734.
- Cooper, A.B., Sawai, C.M., Sicinska, E., Powers, S.E., Sicinski, P., Clark, M.R., and Aifantis, I. (2006). A unique function for cyclin D3 in early B cell development. *Nat. Immunol.* *7*, 489–497.
- Cuylen, S., and Haering, C.H. (2010). A new cohesive team to mediate DNA looping. *Cell Stem Cell* *7*, 424–426.
- D'Addabbo, P., Scascitelli, M., Giambra, V., Rocchi, M., and Frezza, D. (2011). Position and sequence conservation in Amniota of polymorphic enhancer HS1.2 within the palindrome of IgH 3' Regulatory Region. *BMC Evol. Biol.* *11*, 71.
- De Silva, N.S., and Klein, U. (2015). Dynamics of B cells in germinal centres. *Nat. Rev. Immunol.* *15*, 137–148.

- Delpy, L., Le Bert, M., Cogné, M., and Khamlichi, A.A. (2003). Germ-line transcription occurs on both the functional and the non-functional alleles of immunoglobulin constant heavy chain genes. *Eur. J. Immunol.* *33*, 2108–2113.
- Dougier, H.-L., Reynaud, S., Pinaud, E., Carrion, C., Delpy, L., and Cogné, M. (2006). Interallelic class switch recombination can reverse allelic exclusion and allow trans-complementation of an IgH locus switching defect. *Eur. J. Immunol.* *36*, 2181–2191.
- Duquette, M.L., Pham, P., Goodman, M.F., and Maizels, N. (2005). AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation. *Oncogene* *24*, 5791–5798.
- Durandy, A., Kracker, S., and Fischer, A. (2013). Primary antibody deficiencies. *Nat. Rev. Immunol.* *13*, 519–533.
- Early, P., Huang, H., Davis, M., Calame, K., and Hood, L. (2004). An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. 1980. *J. Immunol.* *173*, 6503–6514.
- Ehlich, A., Martin, V., Müller, W., and Rajewsky, K. (1994). Analysis of the B-cell progenitor compartment at the level of single cells. *Curr. Biol.* *4*, 573–583.
- Elmlund, H., Baraznenok, V., Lindahl, M., Samuelsen, C.O., Koeck, P.J.B., Holmberg, S., Hebert, H., and Gustafsson, C.M. (2006). The cyclin-dependent kinase 8 module sterically blocks Mediator interactions with RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.* *103*, 15788–15793.
- Ernst, P., and Smale, S.T. (1995). Combinatorial regulation of transcription II: The immunoglobulin mu heavy chain gene. *Immunity* *2*, 427–438.
- Eyboulet, F., Wydau-Dematteis, S., Eychenne, T., Alibert, O., Neil, H., Boschiero, C., Nevers, M.-C., Volland, H., Cornu, D., Redeker, V., et al. (2015). Mediator independently orchestrates multiple steps of preinitiation complex assembly in vivo. *Nucleic Acids Res.* *43*, 9214–9231.
- Fedderson, R.M., Martin, D.J., and Van Ness, B.G. (1990). The frequency of multiple recombination events occurring at the human Ig kappa L chain locus. *J. Immunol.* *144*, 1088–1093.
- Feldman, S., Achour, I., Wuerffel, R., Kumar, S., Gerasimova, T., Sen, R., and Kenter, A.L. (2015). Constraints contributed by chromatin looping limit recombination targeting during Ig class switch recombination. *J. Immunol.* *194*, 2380–2389.
- Frezza, D., Giambra, V., Cianci, R., Fruscalzo, A., Giufrè, M., Cammarota, G., Martínez-Labarga, C., Rickards, O., Scibilia, G., Sferlazzas, C., et al. (2004). Increased frequency of the immunoglobulin enhancer HS1,2 allele 2 in coeliac disease. *Scand. J. Gastroenterol.* *39*, 1083–1087.
- Frezza, D., Toluoso, B., Giambra, V., Gremese, E., Marchini, M., Nowik, M., Serone, E., D’Addabbo, P., Mattioli, C., Canestri, S., et al. (2012). Polymorphisms of the IgH enhancer HS1.2 and risk of systemic lupus erythematosus. *Ann. Rheum. Dis.* *71*, 1309–1315.
- Gan, W., Guan, Z., Liu, J., Gui, T., Shen, K., Manley, J.L., and Li, X. (2011). R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev.* *25*, 2041–2056.
- Garrett, F.E., Emelyanov, A.V., Sepulveda, M.A., Flanagan, P., Volpi, S., Li, F., Loukinov, D., Eckhardt, L.A., Lobanenko, V.V., and Birshtein, B.K. (2005). Chromatin architecture near a potential 3’ end of the igh locus involves modular regulation of histone modifications during B-Cell development and in vivo occupancy at CTCF sites. *Mol. Cell. Biol.* *25*, 1511–1525.

- Gaudot, L., Thomas-Claudepierre, A.-S., Robert, I., Schiavo, E., Heyer, V., and Reina-San-Martin, B. (2018). Med1 deficiency alters the somatic hypermutation spectrum in murine B cells. *Eur. J. Immunol.* *48*, 720–723.
- Giambra, V., Volpi, S., Emelyanov, A.V., Pflugh, D., Bothwell, A.L.M., Norio, P., Fan, Y., Ju, Z., Skoultschi, A.I., Hardy, R.R., et al. (2008). Pax5 and linker histone H1 coordinate DNA methylation and histone modifications in the 3' regulatory region of the immunoglobulin heavy chain locus. *Mol. Cell Biol.* *28*, 6123–6133.
- Gillies, S.D., Morrison, S.L., Oi, V.T., and Tonegawa, S. (1983). A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* *33*, 717–728.
- Girelli Zubani, G., Zivojnovic, M., De Smet, A., Albagli-Curiel, O., Huetz, F., Weill, J.-C., Reynaud, C.-A., and Storck, S. (2017). Pms2 and uracil-DNA glycosylases act jointly in the mismatch repair pathway to generate Ig gene mutations at A-T base pairs. *J. Exp. Med.* *214*, 1169–1180.
- Giudicelli, V., Chaume, D., and Lefranc, M.-P. (2005). IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res.* *33*, D256-261.
- Gregor, P.D., and Morrison, S.L. (1986). Myeloma mutant with a novel 3' flanking region: loss of normal sequence and insertion of repetitive elements leads to decreased transcription but normal processing of the alpha heavy-chain gene products. *Mol. Cell Biol.* *6*, 1903–1916.
- Hackney, J.A., Misaghi, S., Senger, K., Garris, C., Sun, Y., Lorenzo, M.N., and Zarrin, A.A. (2009). DNA targets of AID evolutionary link between antibody somatic hypermutation and class switch recombination. *Adv. Immunol.* *101*, 163–189.
- Haddad, D., Puget, N., Laviolette-Malirat, N., Conte, C., and Khamlichi, A.A. (2011). Seeking sense of antisense switch transcripts. *Transcription* *2*, 183–188.
- Halverson, R., Torres, R.M., and Pelanda, R. (2004). Receptor editing is the main mechanism of B cell tolerance toward membrane antigens. *Nat. Immunol.* *5*, 645–650.
- Hampel, F., Ehrenberg, S., Hojer, C., Draeseke, A., Marschall-Schröter, G., Kühn, R., Mack, B., Gires, O., Vahl, C.J., Schmidt-Supprian, M., et al. (2011). CD19-independent instruction of murine marginal zone B-cell development by constitutive Notch2 signaling. *Blood* *118*, 6321–6331.
- Han, L., Masani, S., and Yu, K. (2011). Overlapping activation-induced cytidine deaminase hotspot motifs in Ig class-switch recombination. *Proc. Natl. Acad. Sci. U.S.A.* *108*, 11584–11589.
- Hardy, R.R., Kincade, P.W., and Dorshkind, K. (2007). The protean nature of cells in the B lymphocyte lineage. *Immunity* *26*, 703–714.
- Hendriks, R.W., Yuvaraj, S., and Kil, L.P. (2014). Targeting Bruton's tyrosine kinase in B cell malignancies. *Nat. Rev. Cancer* *14*, 219–232.
- Hengartner, C.J., Myer, V.E., Liao, S.M., Wilson, C.J., Koh, S.S., and Young, R.A. (1998). Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol. Cell* *2*, 43–53.
- Hoek, K.L., Antony, P., Lowe, J., Shinnars, N., Sarmah, B., Wenthe, S.R., Wang, D., Gerstein, R.M., and Khan, W.N. (2006). Transitional B cell fate is associated with developmental stage-specific regulation of diacylglycerol and calcium signaling upon B cell receptor engagement. *J. Immunol.* *177*, 5405–5413.

- Hoffman, W., Lakkis, F.G., and Chalasani, G. (2016). B Cells, Antibodies, and More. *Clin J Am Soc Nephrol* 11, 137–154.
- Hwang, J.K., Alt, F.W., and Yeap, L.-S. (2015). Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. *Microbiol Spectr* 3, MDNA3-0037–2014.
- Jung, D., Giallourakis, C., Mostoslavsky, R., and Alt, F.W. (2006). Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu. Rev. Immunol.* 24, 541–570.
- Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., Berkum, N.L. van, Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430–435.
- Kanayama, N., Cascalho, M., and Ohmori, H. (2005). Analysis of marginal zone B cell development in the mouse with limited B cell diversity: role of the antigen receptor signals in the recruitment of B cells to the marginal zone. *J. Immunol.* 174, 1438–1445.
- Kassambara, A., Jourdan, M., Bruyer, A., Robert, N., Pantesco, V., Elemento, O., Klein, B., and Moreaux, J. (2017). Global miRNA expression analysis identifies novel key regulators of plasma cell differentiation and malignant plasma cell. *Nucleic Acids Res* 45, 5639–5652.
- Kaufmann, R., Straussberg, R., Mandel, H., Fattal-Valevski, A., Ben-Zeev, B., Naamati, A., Shaag, A., Zenvirt, S., Konen, O., Mimouni-Bloch, A., et al. (2010). Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. *Am. J. Hum. Genet.* 87, 667–670.
- Kelleher, R.J., Flanagan, P.M., and Kornberg, R.D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* 61, 1209–1215.
- Kenter, A.L., Wuerffel, R., Kumar, S., and Grigera, F. (2013). Genomic Architecture may Influence Recurrent Chromosomal Translocation Frequency in the Igh Locus. *Front Immunol* 4, 500.
- Khamlichi, A.A., Pinaud, E., Decourt, C., Chauveau, C., and Cogné, M. (2000). The 3' IgH regulatory region: a complex structure in a search for a function. *Adv. Immunol.* 75, 317–345.
- Khamlichi, A.A., Glaudet, F., Oruc, Z., Denis, V., Le Bert, M., and Cogné, M. (2004). Immunoglobulin class-switch recombination in mice devoid of any S mu tandem repeat. *Blood* 103, 3828–3836.
- Kim, Y.J., Björklund, S., Li, Y., Sayre, M.H., and Kornberg, R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* 77, 599–608.
- Kingzette, M., Spieker-Polet, H., Yam, P.C., Zhai, S.K., and Knight, K.L. (1998). Trans-chromosomal recombination within the Ig heavy chain switch region in B lymphocytes. *Proc. Natl. Acad. Sci. U.S.A.* 95, 11840–11845.
- Klein, U., Rajewsky, K., and Küppers, R. (1998). Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. *J. Exp. Med.* 188, 1679–1689.
- Knuesel, M.T., Meyer, K.D., Bernecky, C., and Taatjes, D.J. (2009). The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev.* 23, 439–451.

- Kotnis, A., Du, L., Liu, C., Popov, S.W., and Pan-Hammarström, Q. (2009). Non-homologous end joining in class switch recombination: the beginning of the end. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* *364*, 653–665.
- Kurosaki, T., Kometani, K., and Ise, W. (2015). Memory B cells. *Nat. Rev. Immunol.* *15*, 149–159.
- Laffleur, B., Denis-Lagache, N., Péron, S., Sirac, C., Moreau, J., and Cogné, M. (2014a). AID-induced remodeling of immunoglobulin genes and B cell fate. *Oncotarget* *5*, 1118–1131.
- Laffleur, B., Bardet, S.M., Garot, A., Brousse, M., Baylet, A., and Cogné, M. (2014b). Immunoglobulin genes undergo legitimate repair in human B cells not only after cis- but also frequent trans-class switch recombination. *Genes Immun.* *15*, 341–346.
- Le Noir, S., Boyer, F., Lecardeur, S., Brousse, M., Oruc, Z., Cook-Moreau, J., Denizot, Y., and Cogné, M. (2017a). Functional anatomy of the immunoglobulin heavy chain 3' super-enhancer needs not only core enhancer elements but also their unique DNA context. *Nucleic Acids Res.* *45*, 5829–5837.
- Le Noir, S., Laffleur, B., Carrion, C., Garot, A., Lecardeur, S., Pinaud, E., Denizot, Y., Skok, J., and Cogné, M. (2017b). The IgH locus 3' cis-regulatory super-enhancer co-opts AID for allelic transvection. *Oncotarget*.
- Li, J., and Xu, X. (2016). DNA double-strand break repair: a tale of pathway choices. *Acta Biochim. Biophys. Sin. (Shanghai)* *48*, 641–646.
- Li, G., White, C.A., Lam, T., Pone, E.J., Tran, D.C., Hayama, K.L., Zan, H., Xu, Z., and Casali, P. (2013). Combinatorial H3K9acS10ph Histone Modifications in IgH Locus S Regions Target 14-3-3 Adaptors and AID to Specify Antibody Class Switch DNA Recombination. *Cell Rep* *5*, 702–714.
- Lieberson, R., Ong, J., Shi, X., and Eckhardt, L.A. (1995). Immunoglobulin gene transcription ceases upon deletion of a distant enhancer. *EMBO J.* *14*, 6229–6238.
- Liu, Y.J., Malisan, F., de Bouteiller, O., Guret, C., Lebecque, S., Banchereau, J., Mills, F.C., Max, E.E., and Martinez-Valdez, H. (1996). Within germinal centers, isotype switching of immunoglobulin genes occurs after the onset of somatic mutation. *Immunity* *4*, 241–250.
- Liubchenko, G.A., Appleberry, H.C., Holers, V.M., Banda, N.K., Willis, V.C., and Lyubchenko, T. (2012). Potentially autoreactive naturally occurring transitional T3 B lymphocytes exhibit a unique signaling profile. *J. Autoimmun.* *38*, 293–303.
- Macias-Garcia, A., Heizmann, B., Sellars, M., Marchal, P., Dali, H., Pasquali, J.-L., Muller, S., Kastner, P., and Chan, S. (2016). Ikaros Is a Negative Regulator of B1 Cell Development and Function. *J. Biol. Chem.* *291*, 9073–9086.
- Manis, J.P., van der Stoep, N., Tian, M., Ferrini, R., Davidson, L., Bottaro, A., and Alt, F.W. (1998). Class switching in B cells lacking 3' immunoglobulin heavy chain enhancers. *J. Exp. Med.* *188*, 1421–1431.
- Marie-Cardine, A., Divay, F., Dutot, I., Green, A., Perdrix, A., Boyer, O., Contentin, N., Tilly, H., Tron, F., Vannier, J.-P., et al. (2008). Transitional B cells in humans: characterization and insight from B lymphocyte reconstitution after hematopoietic stem cell transplantation. *Clin. Immunol.* *127*, 14–25.
- Marquet, M., Garot, A., Bender, S., Carrion, C., Rouaud, P., Lecardeur, S., Denizot, Y., Cogné, M., and Pinaud, E. (2014). The E μ enhancer region influences H chain expression and B cell fate without impacting IgVH repertoire and immune response in vivo. *J. Immunol.* *193*, 1171–1183.

- Mårtensson, I.-L., Almqvist, N., Grimsholm, O., and Bernardi, A.I. (2010). The pre-B cell receptor checkpoint. *FEBS Letters* 584, 2572–2579.
- Martin, V.G., Wu, Y.-C.B., Townsend, C.L., Lu, G.H.C., O’Hare, J.S., Mozeika, A., Coolen, A.C.C., Kipling, D., Fraternali, F., and Dunn-Walters, D.K. (2016). Transitional B Cells in Early Human B Cell Development – Time to Revisit the Paradigm? *Front Immunol* 7.
- Matsuoka, M., Yoshida, K., Maeda, T., Usuda, S., and Sakano, H. (1990). Switch circular DNA formed in cytokine-treated mouse splenocytes: evidence for intramolecular DNA deletion in immunoglobulin class switching. *Cell* 62, 135–142.
- Mehra, P., Gerasimova, T., Basu, A., Jha, V., Banerjee, A., Sindhava, V., Gray, F., Berry, C.T., Sen, R., and Atchison, M.L. (2016). YY1 controls E μ -3’RR DNA loop formation and immunoglobulin heavy chain class switch recombination. *Blood Adv* 1, 15–20.
- Michaelson, J.S., Giannini, S.L., and Birshtein, B.K. (1995). Identification of 3 α -hs4, a novel Ig heavy chain enhancer element regulated at multiple stages of B cell differentiation. *Nucleic Acids Res* 23, 975–981.
- Mourcin, F., Pangault, C., Amin-Ali, R., Amé-Thomas, P., and Tarte, K. (2012). Stromal cell contribution to human follicular lymphoma pathogenesis. *Front Immunol* 3, 280.
- Mukundan, B., and Ansari, A. (2011). Novel role for mediator complex subunit Srb5/Med18 in termination of transcription. *J. Biol. Chem.* 286, 37053–37057.
- Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O., and 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. *J. Biol. Chem.* 274, 18470–18476.
- Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. *Cell* 102, 553–563.
- Nagasawa, T. (2006). Microenvironmental niches in the bone marrow required for B-cell development. *Nat. Rev. Immunol.* 6, 107–116.
- Napoli, C., Sessa, M., Infante, T., and Casamassimi, A. (2012). Unraveling framework of the ancestral Mediator complex in human diseases. *Biochimie* 94, 579–587.
- Nemazee, D. (2017). Mechanisms of central tolerance for B cells. *Nat Rev Immunol* 17, 281–294.
- Neumann, B., Klippert, A., Raue, K., Sopper, S., and Stahl-Hennig, C. (2015). Characterization of B and plasma cells in blood, bone marrow, and secondary lymphoid organs of rhesus macaques by multicolor flow cytometry. *J. Leukoc. Biol.* 97, 19–30.
- Nicolas, L., Cols, M., Choi, J.E., Chaudhuri, J., and Vuong, B. (2018). Generating and repairing genetically programmed DNA breaks during immunoglobulin class switch recombination. *F1000Res* 7, 458.
- Niuro, H., and Clark, E.A. (2002). Regulation of B-cell fate by antigen-receptor signals. *Nat. Rev. Immunol.* 2, 945–956.
- Nutt, S.L., Hodgkin, P.D., Tarlinton, D.M., and Corcoran, L.M. (2015). The generation of antibody-secreting plasma cells. *Nat. Rev. Immunol.* 15, 160–171.

- Odegard, V.H., and Schatz, D.G. (2006). Targeting of somatic hypermutation. *Nat. Rev. Immunol.* 6, 573–583.
- Oliveira, F.L., Aguiar, A.M., Borojevic, R., and El-Cheikh, M.C. (2005). IgE expression on the surface of B1 and B2 lymphocytes in experimental murine schistosomiasis. *Braz. J. Med. Biol. Res.* 38, 1033–1042.
- Oracki, S.A., Walker, J.A., Hibbs, M.L., Corcoran, L.M., and Tarlinton, D.M. (2010). Plasma cell development and survival. *Immunol. Rev.* 237, 140–159.
- Pan-Hammarström, Q., Jones, A.-M., Lähdesmäki, A., Zhou, W., Gatti, R.A., Hammarström, L., Gennery, A.R., and Ehrenstein, M.R. (2005). Impact of DNA ligase IV on nonhomologous end joining pathways during class switch recombination in human cells. *J. Exp. Med.* 201, 189–194.
- Patenaude, A.-M., and Di Noia, J.M. (2010). The mechanisms regulating the subcellular localization of AID. *Nucleus* 1, 325–331.
- Pavri, R. (2017). R Loops in the Regulation of Antibody Gene Diversification. *Genes (Basel)* 8.
- Pelayo, R., Dorantes-Acosta, E., Vadillo, E., and Fuentes-Panana, E. (2012). From HSC to B-Lymphoid Cells in Normal and Malignant Hematopoiesis. *Advances in Hematopoietic Stem Cell Research*.
- Peled, J.U., Kuang, F.L., Iglesias-Ussel, M.D., Roa, S., Kalis, S.L., Goodman, M.F., and Scharff, M.D. (2008). The biochemistry of somatic hypermutation. *Annu. Rev. Immunol.* 26, 481–511.
- Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A., and Bejerano, G. (2013). Enhancers: five essential questions. *Nat. Rev. Genet.* 14, 288–295.
- Péron, S., Laffleur, B., Denis-Lagache, N., Cook-Moreau, J., Tinguely, A., Delpy, L., Denizot, Y., Pinaud, E., and Cogné, M. (2012a). AID-driven deletion causes immunoglobulin heavy chain locus suicide recombination in B cells. *Science* 336, 931–934.
- Péron, S., Laffleur, B., Denis-Lagache, N., Cook-Moreau, J., Filloux, M., and Cogné, M. (2012b). [IgH locus suicide recombination, or when B cells surrender!]. *Med Sci (Paris)* 28, 551–554.
- Petro, J.B., Gerstein, R.M., Lowe, J., Carter, R.S., Shinnars, N., and Khan, W.N. (2002). Transitional type 1 and 2 B lymphocyte subsets are differentially responsive to antigen receptor signaling. *J. Biol. Chem.* 277, 48009–48019.
- Pieper, K., Grimbacher, B., and Eibel, H. (2013). B-cell biology and development. *J. Allergy Clin. Immunol.* 131, 959–971.
- Pillai, S., and Cariappa, A. (2009a). The follicular versus marginal zone B lymphocyte cell fate decision. *Nat. Rev. Immunol.* 9, 767–777.
- Pillai, S., and Cariappa, A. (2009b). The bone marrow perisinusoidal niche for recirculating B cells and the positive selection of bone marrow-derived B lymphocytes. *Immunol. Cell Biol.* 87, 16–19.
- Pinaud, E., Aupetit, C., Chauveau, C., and Cogné, M. (1997). Identification of a homolog of the C alpha 3'/hs3 enhancer and of an allelic variant of the 3'IgH/hs1,2 enhancer downstream of the human immunoglobulin alpha 1 gene. *Eur. J. Immunol.* 27, 2981–2985.
- Pinaud, E., Khamlichi, A.A., Le Morvan, C., Drouet, M., Nalesso, V., Le Bert, M., and Cogné, M. (2001). Localization of the 3' IgH locus elements that effect long-distance regulation of class switch recombination. *Immunity* 15, 187–199.

- Pinaud, E., Marquet, M., Fiancette, R., Péron, S., Vincent-Fabert, C., Denizot, Y., and Cogné, M. (2011). The IgH locus 3' regulatory region: pulling the strings from behind. *Adv. Immunol.* *110*, 27–70.
- Pone, E.J., Zhang, J., Mai, T., White, C.A., Li, G., Sakakura, J.K., Patel, P.J., Al-Qahtani, A., Zan, H., Xu, Z., et al. (2012). BCR-signalling synergizes with TLR-signalling for induction of AID and immunoglobulin class-switching through the non-canonical NF- κ B pathway. *Nat Commun* *3*, 767.
- Rada, C., Williams, G.T., Nilsen, H., Barnes, D.E., Lindahl, T., and Neuberger, M.S. (2002). Immunoglobulin Isotype Switching Is Inhibited and Somatic Hypermutation Perturbed in UNG-Deficient Mice. *Current Biology* *12*, 1748–1755.
- Rassenti, L.Z., Pratt, L.F., Chen, P.P., Carson, D.A., and Kipps, T.J. (1991). Autoantibody-encoding kappa L chain genes frequently rearranged in lambda L chain-expressing chronic lymphocytic leukemia. *J. Immunol.* *147*, 1060–1066.
- Rauta, P.R., Nayak, B., and Das, S. (2012). Immune system and immune responses in fish and their role in comparative immunity study: A model for higher organisms. *Immunology Letters* *148*, 23–33.
- Rebhandl, S., Huemer, M., Greil, R., and Geisberger, R. (2015). AID/APOBEC deaminases and cancer. *Oncoscience* *2*, 320–333.
- Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labeau, R., Gennery, A., et al. (2000). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* *102*, 565–575.
- Reynaud, C.-A., Delbos, F., Faily, A., Gueranger, Q., Aoufouchi, S., and Weill, J.-C. (2009). Competitive repair pathways in immunoglobulin gene hypermutation. *Philosophical Transactions of the Royal Society B: Biological Sciences* *364*, 613–619.
- Reynaud, C.-A., Descatoire, M., Dogan, I., Huetz, F., Weller, S., and Weill, J.-C. (2012). IgM memory B cells: a mouse/human paradox. *Cell. Mol. Life Sci.* *69*, 1625–1634.
- Reynaud, S., Delpy, L., Fleury, L., Dougier, H.-L., Sirac, C., and Cogné, M. (2005). Interallelic class switch recombination contributes significantly to class switching in mouse B cells. *J. Immunol.* *174*, 6176–6183.
- Rhee, H.S., and Pugh, B.F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* *483*, 295–301.
- Rhodes, D., and Lipps, H.J. (2015). G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res* *43*, 8627–8637.
- Ribeiro de Almeida, C., Dhir, S., Dhir, A., Moghaddam, A.E., Sattentau, Q., Meinhart, A., and Proudfoot, N.J. (2018). RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Mol. Cell* *70*, 650-662.e8.
- Riether, C., Schürch, C.M., and Ochsenein, A.F. (2015). Regulation of hematopoietic and leukemic stem cells by the immune system. *Cell Death Differ.* *22*, 187–198.
- Rogozin, I.B., and Diaz, M. (2004). Cutting Edge: DGYW/WRCH Is a Better Predictor of Mutability at G:C Bases in Ig Hypermutation Than the Widely Accepted RGYW/WRCY Motif and Probably Reflects a Two-Step Activation-Induced Cytidine Deaminase-Triggered Process. *The Journal of Immunology* *172*, 3382–3384.

- Roth, D.B. (2003). Restraining the V(D)J recombinase. *Nat. Rev. Immunol.* 3, 656–666.
- Rothman, P., Chen, Y.Y., Lutzker, S., Li, S.C., Stewart, V., Coffman, R., and Alt, F.W. (1990). Structure and expression of germ line immunoglobulin heavy-chain epsilon transcripts: interleukin-4 plus lipopolysaccharide-directed switching to C epsilon. *Mol. Cell. Biol.* 10, 1672–1679.
- Rouaud, P., Vincent-Fabert, C., Fiancette, R., Cogné, M., Pinaud, E., and Denizot, Y. (2012). Enhancers located in heavy chain regulatory region (hs3a, hs1,2, hs3b, and hs4) are dispensable for diversity of VDJ recombination. *J. Biol. Chem.* 287, 8356–8360.
- Rouaud, P., Vincent-Fabert, C., Saintamand, A., Fiancette, R., Marquet, M., Robert, I., Reina-San-Martin, B., Pinaud, E., Cogné, M., and Denizot, Y. (2013). The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. *J. Exp. Med.* 210, 1501–1507.
- Sacquin, A., Gador, M., and Fazilleau, N. (2017). The strength of BCR signaling shapes terminal development of follicular helper T cells in mice. *Eur. J. Immunol.* 47, 1295–1304.
- Saintamand, A., Vincent-Fabert, C., Garot, A., Rouaud, P., Oruc, Z., Magnone, V., Cogné, M., and Denizot, Y. (2016). Deciphering the importance of the palindromic architecture of the immunoglobulin heavy-chain 3' regulatory region. *Nat Commun* 7, 10730.
- Saintamand, A., Vincent-Fabert, C., Marquet, M., Ghazzoui, N., Magnone, V., Pinaud, E., Cogné, M., and Denizot, Y. (2017). $\epsilon\mu$ and 3'RR IgH enhancers show hierarchic unilateral dependence in mature B-cells. *Sci Rep* 7.
- Saito, T., Chiba, S., Ichikawa, M., Kunisato, A., Asai, T., Shimizu, K., Yamaguchi, T., Yamamoto, G., Seo, S., Kumano, K., et al. (2003). Notch2 is preferentially expressed in mature B cells and indispensable for marginal zone B lineage development. *Immunity* 18, 675–685.
- Schlissel, M.S. (2004). Regulation of activation and recombination of the murine Igkappa locus. *Immunol. Rev.* 200, 215–223.
- Schroeder, H.W., and Cavacini, L. (2010). Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* 125, S41-52.
- von Schwedler, U., Jäck, H.M., and Wabl, M. (1990). Circular DNA is a product of the immunoglobulin class switch rearrangement. *Nature* 345, 452–456.
- Shapiro-Shelef, M., and Calame, K. (2005). Regulation of plasma-cell development. *Nat. Rev. Immunol.* 5, 230–242.
- Shetty, K., and Schatz, D.G. (2015). Recruitment of RAG1 and RAG2 to Chromatinized DNA during V(D)J Recombination. *Mol. Cell. Biol.* 35, 3701–3713.
- Shinkura, R., Tian, M., Smith, M., Chua, K., Fujiwara, Y., and Alt, F.W. (2003). The influence of transcriptional orientation on endogenous switch region function. *Nat. Immunol.* 4, 435–441.
- Shinkura, R., Ito, S., Begum, N.A., Nagaoka, H., Muramatsu, M., Kinoshita, K., Sakakibara, Y., Hijikata, H., and Honjo, T. (2004). Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat. Immunol.* 5, 707–712.
- Siminovitch, K.A., Bakhshi, A., Goldman, P., and Korsmeyer, S.J. (1985). A uniform deleting element mediates the loss of kappa genes in human B cells. *Nature* 316, 260–262.

- Soulas-Sprauel, P., Le Guyader, G., Rivera-Munoz, P., Abramowski, V., Olivier-Martin, C., Goujet-Zalc, C., Charneau, P., and de Villartay, J.-P. (2007). Role for DNA repair factor XRCC4 in immunoglobulin class switch recombination. *J. Exp. Med.* *204*, 1717–1727.
- Soutourina, J. (2018). Transcription regulation by the Mediator complex. *Nat. Rev. Mol. Cell Biol.* *19*, 262–274.
- Soutourina, J., Wydau, S., Ambroise, Y., Boschiero, C., and Werner, M. (2011). Direct interaction of RNA polymerase II and mediator required for transcription in vivo. *Science* *331*, 1451–1454.
- Spitler, K.M., Ponce, J.M., Oudit, G.Y., Hall, D.D., and Grueter, C.E. (2017). Cardiac Med1 deletion promotes early lethality, cardiac remodeling, and transcriptional reprogramming. *Am. J. Physiol. Heart Circ. Physiol.* *312*, H768–H780.
- Stavnezer, J., Björkman, A., Du, L., Cagigi, A., and Pan-Hammarström, Q. (2010). Mapping of switch recombination junctions, a tool for studying DNA repair pathways during immunoglobulin class switching. *Adv. Immunol.* *108*, 45–109.
- Steele, A.J. (2014). PI3K in CLL: are 2 isoforms better than 1? *Blood* *124*, 3508–3510.
- Stevenson, F.K., Krysov, S., Davies, A.J., Steele, A.J., and Packham, G. (2011). B-cell receptor signaling in chronic lymphocytic leukemia. *Blood* *118*, 4313–4320.
- Streubel, G., and Bracken, A.P. (2015). MED23: a new Mediator of H2B monoubiquitylation. *EMBO J.* *34*, 2863–2864.
- Takahashi, H., Parmely, T.J., Sato, S., Tomomori-Sato, C., Banks, C.A.S., Kong, S.E., Szutorisz, H., Swanson, S.K., Martin-Brown, S., Washburn, M.P., et al. (2011). Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* *146*, 92–104.
- Tangye, S.G., Ma, C.S., Brink, R., and Deenick, E.K. (2013). The good, the bad and the ugly - TFH cells in human health and disease. *Nat. Rev. Immunol.* *13*, 412–426.
- Teague, B.N., Pan, Y., Mudd, P.A., Nakken, B., Zhang, Q., Szodoray, P., Kim-Howard, X., Wilson, P.C., and Farris, A.D. (2007). Cutting edge: Transitional T3 B cells do not give rise to mature B cells, have undergone selection, and are reduced in murine lupus. *J. Immunol.* *178*, 7511–7515.
- Tellier, J., Shi, W., Minnich, M., Liao, Y., Crawford, S., Smyth, G.K., Kallies, A., Busslinger, M., and Nutt, S.L. (2016). Blimp-1 controls plasma cell function through regulation of immunoglobulin secretion and the unfolded protein response. *Nat Immunol* *17*, 323–330.
- Thai, L.-H., Le Gallou, S., Robbins, A., Crickx, E., Fadeev, T., Zhou, Z., Cagnard, N., Mégret, J., Bole, C., Weill, J.-C., et al. (2018). BAFF and CD4+ T cells are major survival factors for long-lived splenic plasma cells in a B-cell-depletion context. *Blood* *131*, 1545–1555.
- Thomas-Claudepierre, A.-S., Schiavo, E., Heyer, V., Fournier, M., Page, A., Robert, I., and Reina-San-Martin, B. (2013). The cohesin complex regulates immunoglobulin class switch recombination. *J. Exp. Med.* *210*, 2495–2502.
- 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., et al. (2016). Mediator facilitates transcriptional activation and dynamic long-range contacts at the IgH locus during class switch recombination. *Journal of Experimental Medicine* *213*, 303–312.

- Tsai, K.-L., Sato, S., Tomomori-Sato, C., Conaway, R.C., Conaway, J.W., and Asturias, F.J. (2013). A conserved Mediator-CDK8 kinase module association regulates Mediator-RNA polymerase II interaction. *Nat. Struct. Mol. Biol.* *20*, 611–619.
- Tsai, K.-L., Tomomori-Sato, C., Sato, S., Conaway, R.C., Conaway, J.W., and Asturias, F.J. (2014). Subunit architecture and functional modular rearrangements of the transcriptional mediator complex. *Cell* *157*, 1430–1444.
- Tumang, J.R., Holodick, N.E., Vizconde, T.C., Kaku, H., Frances, R., and Rothstein, T.L. (2011). A CD25⁺ Positive Population of Activated B1 Cells Expresses LIFR and Responds to LIF. *Front. Immunol.* *2*.
- Tümekaya, T., van der Burg, M., Garcia Sanz, R., Gonzalez Diaz, M., Langerak, A.W., San Miguel, J.F., and van Dongen, J.J. (2001). Immunoglobulin lambda isotype gene rearrangements in B cell malignancies. *Leukemia* *15*, 121–127.
- Vaidyanathan, B., Yen, W.-F., Pucella, J.N., and Chaudhuri, J. (2014). AIDing Chromatin and Transcription-Coupled Orchestration of Immunoglobulin Class-Switch Recombination. *Front Immunol* *5*, 120.
- Vincent-Fabert, C., Truffinet, V., Fiancette, R., Cogné, N., Cogné, M., and Denizot, Y. (2009). Ig synthesis and class switching do not require the presence of the hs4 enhancer in the 3' IgH regulatory region. *J. Immunol.* *182*, 6926–6932.
- Vincent-Fabert, C., Fiancette, R., Pinaud, E., Truffinet, V., Cogné, N., Cogné, M., and Denizot, Y. (2010). Genomic deletion of the whole IgH 3' regulatory region (hs3a, hs1,2, hs3b, and hs4) dramatically affects class switch recombination and Ig secretion to all isotypes. *Blood* *116*, 1895–1898.
- Volpi, S.A., Verma-Gaur, J., Hassan, R., Ju, Z., Roa, S., Chatterjee, S., Werling, U., Hou, H., Will, B., Steidl, U., et al. (2012). Germline deletion of Igh 3' regulatory region elements hs5-7 affects B cell specific regulation, rearrangement and insulation of the Igh locus. *J Immunol* *188*, 2556–2566.
- Wang, W., Yao, X., Huang, Y., Hu, X., Liu, R., Hou, D., Chen, R., and Wang, G. (2013). Mediator MED23 regulates basal transcription in vivo via an interaction with P-TEFb. *Transcription* *4*, 39–51.
- Weigert, M.G., Cesari, I.M., Yonkovich, S.J., and Cohn, M. (1970). Variability in the Lambda Light Chain Sequences of Mouse Antibody. *Nature* *228*, 1045–1047.
- Witt, C.M., Won, W.-J., Hurez, V., and Klug, C.A. (2003). Notch2 haploinsufficiency results in diminished B1 B cells and a severe reduction in marginal zone B cells. *J. Immunol.* *171*, 2783–2788.
- Wong, K.H., Jin, Y., and Struhl, K. (2014). TFIIH phosphorylation of the Pol II CTD stimulates mediator dissociation from the preinitiation complex and promoter escape. *Mol. Cell* *54*, 601–612.
- Wuerffel, R., Wang, L., Grigera, F., Manis, J., Selsing, E., Perlot, T., Alt, F.W., Cogne, M., Pinaud, E., and Kenter, A.L. (2007). S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity* *27*, 711–722.
- Xu, Z., Pone, E.J., Al-Qahtani, A., Park, S.-R., Zan, H., and Casali, P. (2007). Regulation of aicda expression and AID activity: relevance to somatic hypermutation and class switch DNA recombination. *Crit. Rev. Immunol.* *27*, 367–397.
- Xu, Z., Zan, H., Pone, E.J., Mai, T., and Casali, P. (2012). Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat. Rev. Immunol.* *12*, 517–531.

- Yaari, G., Benichou, J.I.C., Vander Heiden, J.A., Kleinstein, S.H., and Louzoun, Y. (2015). The mutation patterns in B-cell immunoglobulin receptors reflect the influence of selection acting at multiple time-scales. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370, 20140242.
- Yan, C.T., Boboila, C., Souza, E.K., Franco, S., Hickernell, T.R., Murphy, M., Gumaste, S., Geyer, M., Zarrin, A.A., Manis, J.P., et al. (2007). IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* 449, 478–482.
- Yin, J., and Wang, G. (2014). The Mediator complex: a master coordinator of transcription and cell lineage development. *Development* 141, 977–987.
- Youn, D.Y., Xiaoli, A.M., Pessin, J.E., and Yang, F. (2016). Regulation of metabolism by the Mediator complex. *Biophys Rep* 2, 69–77.
- Yusuf, I., Zhu, X., Kharas, M.G., Chen, J., and Fruman, D.A. (2004). Optimal B-cell proliferation requires phosphoinositide 3-kinase-dependent inactivation of FOXO transcription factors. *Blood* 104, 784–787.
- Zan, H., and Casali, P. (2015). Epigenetics of Peripheral B-Cell Differentiation and the Antibody Response. *Front Immunol* 6.
- Zarrin, A.A., Alt, F.W., Chaudhuri, J., Stokes, N., Kaushal, D., Du Pasquier, L., and Tian, M. (2004). An evolutionarily conserved target motif for immunoglobulin class-switch recombination. *Nat. Immunol.* 5, 1275–1281.
- Zhang, M., Srivastava, G., and Lu, L. (2004). The pre-B cell receptor and its function during B cell development. *Cell. Mol. Immunol.* 1, 89–94.
- Zhang, Q., Iida, R., Yokota, T., and Kincade, P.W. (2013). Early events in lymphopoiesis: an update. *Curr. Opin. Hematol.* 20, 265–272.
- Zhang, Z.Z., Pannunzio, N.R., Hsieh, C.-L., Yu, K., and Lieber, M.R. (2014). The role of G-density in switch region repeats for immunoglobulin class switch recombination. *Nucleic Acids Res.* 42, 13186–13193.
- Zheng, S., Vuong, B.Q., Vaidyanathan, B., Lin, J.-Y., Huang, F.-T., and Chaudhuri, J. (2015). Non-coding RNA Generated following Lariat Debranching Mediates Targeting of AID to DNA. *Cell* 161, 762–773.
- Zouali, M., and Richard, Y. (2011). Marginal zone B-cells, a gatekeeper of innate immunity. *Front Immunol* 2, 63.

Résumé

L'activation des cellules B est connue pour s'accompagner de remodelages des gènes d'immunoglobulines qui permettent la maturation d'affinité des régions variables d'Ig par hypermutation somatique SHM et la commutation de classe CSR. Ces deux processus sont sous le contrôle de la région régulatrice 3' (3'RR) du locus IgH. Pendant la CSR, le locus IgH subit des changements tridimensionnels mettant les régions switch ciblées par AID à proximité de la région 3'RR afin de faciliter la recombinaison. La sous-unité MED1 du complexe Médiateur favorise cette interaction à longue distance avec la 3'RR mais elle intervient aussi dans la transcription germinale qui précède la CSR afin de faciliter l'activité d'AID. Comme récemment démontré chez la souris, la région 3'RR peut aussi être ciblée par des recombinaison médiée par AID, mais contrairement à la CSR, ce type de recombinaison qui joint la région S μ et la 3'RR et qui s'appelle recombinaison suicide du locus IgH ou LSR entraîne une délétion complète de l'ensemble des gènes constants conduisant à la mort des cellules B par la perte de l'expression du BCR. Nous montrons maintenant que la LSR induite par AID se produit également dans les cellules B humaines activées avec les deux régions 3'RR (3'RR1 en aval de C α 1 et 3'RR2 en aval de C α 2) et elle touche l'allèle fonctionnel. Elle peut aussi être bi-allélique marquée par l'absence de ce type de recombinaison dans les plasmocytes de la moelle mais aussi dans les cellules B mémoires au repos du sang. La LSR se produit à un niveau élevé lorsque les cellules B mémoires sont réactivées. Toutes nos conditions de stimulation utilisées *in-vitro* induit la LSR, sans permettre de discerner comment se fait « le choix » entre la CSR et la LSR. Nos résultats montrent par contre que la sous-unité MED1 semble influencer la transcription de la 3'RR et la recombinaison LSR chez la souris. L'inactivation conditionnelle de MED1 influence l'accessibilité transcriptionnelle et donc les recombinaisons sans affecter les marques épigénétiques du locus IgH. Cette étude de MED1 a aussi révélé que l'ensemble des processus stimulés par l'IgH 3'RR sont « Médiateur-dépendants » (SHM, CSR sans distinction de la cis et la trans-CSR, expression augmentée du locus dans les plasmocytes...), comme semble l'être également le processus de choix des segments variables au cours des réarrangements V_HDJ_H.

Mots-clés : 3'RR, locus IgH, CSR, LSR, complexe Médiateur.

Summary

B-cell activation is accompanied by remodeling of immunoglobulin genes resulting in affinity maturation of Ig variable regions by somatic hypermutation (SHM) and class switch recombination (CSR). These two processes are under the control of the 3' regulatory region (3'RR) of the IgH locus. During CSR, the IgH locus undergoes three dimensional changes bringing the AID-targeted switch regions near the 3'RR region to facilitate recombination. The MED1 subunit of the Mediator complex promotes this long-distance interaction with the 3'RR, but it is also implicated in germinal transcription preceding CSR in order to facilitate AID activity. As recently demonstrated in mice, the 3'RR region can also be targeted by AID-mediated recombination, but unlike CSR, this type of recombination joining the S μ region and 3'RR (called Locus Suicide Recombination or LSR) results in a complete deletion of all the constant genes leading to B-cell death by loss of B Cell Receptor expression. We now show that AID-mediated LSR also occurs in activated human B cells with the two 3'RR (3'RR1 downstream of C α 1 and 3'RR2 downstream of C α 2) and affects the functional allele. It can also be bi-allelic marked by the absence of this type of recombination in plasma cells of the bone marrow but also in quiescent blood memory B cells. LSR occurs at high level when the memory B cells are reactivated. All *in-vitro* stimulations induce LSR, without identifying conditions favoring either CSR and the LSR. Our results also show that the MED1 subunit appears to influence 3' RR transcription and LSR in mice. Conditional inactivation of MED1 influences transcriptional accessibility and therefore recombination without affecting epigenetic markers of the IgH locus. This study also revealed that all the processes controlled by the 3'RR are "mediator-dependent" (SHM, CSR without distinction between *Cis* and *Trans*-CSR, increased expression of the IgH locus in the plasma cells ...), as well as the choice of variable segments during V_HDJ_H rearrangements.

Keywords: 3'RR, IgH locus, CSR, LSR Mediator complex

