

# frontiers

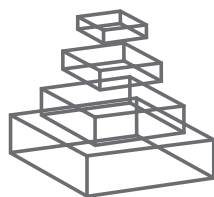
## RESEARCH TOPICS

### CHROMATIN & TRANSCRIPTIONAL TANGO ON THE IMMUNE DANCE FLOOR

Topic Editor  
Ananda L. Roy



frontiers in  
**IMMUNOLOGY**



# frontiers

## FRONTIERS COPYRIGHT STATEMENT

© Copyright 2007-2015  
Frontiers Media SA.  
All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-510-7

DOI 10.3389/978-2-88919-510-7

## ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# CHROMATIN & TRANSCRIPTIONAL TANGO ON THE IMMUNE DANCE FLOOR

Topic Editor:

**Ananda L. Roy**, Tufts University School of Medicine, Boston, USA

Signaling through the cell surface antigen receptor is a hallmark of various stages of lymphocyte development and adaptive immunity. Besides the adaptive immune system, the innate immunity is equally important for protection. However, the mechanistic connection between signaling, chromatin changes and downstream transcriptional pathways in both innate and adaptive immune system remains incompletely understood in hematopoiesis. A related issue is how the enhancers communicate to the promoters in a stage specific fashion and in the context of chromatin. Because the factors that regulate chromatin are generally present and active in most cell types, how could cell type and/or stage specific chromatin architecture be achieved in response to a particular immune signal?

The genetic loci that encode lymphocyte cell surface receptors are in an ‘unrearranged’ or “germline” configuration during the early stages of development. Thus, in addition to expressing lineage and/or stage specific transcription factors during each developmental stage, lymphocytes also need to rearrange their cognate receptor loci in a strictly ordered fashion. Hence, there must be a tightly coordinated communication between the recombination machinery and the transcriptional machinery (including chromatin regulators) at every developmental step. Mature B cells also undergo class-switch recombination and somatic hypermutation. Importantly, along the way, these cells must avoid autoimmune responses and only those cells capable of recognizing foreign-antigens are preserved to reach peripheral organs where they must function. The exquisite regulation that govern chromatin accessibility, recombination and transcription regulation in response to the environmental signals in the immune system is discussed here in a series of articles.

# Table of Contents

<b>05</b>	<b><i>Chromatin and transcriptional tango on the immune dance floor</i></b>	Ananda L. Roy and Robert G. Roeder
<b>08</b>	<b><i>A new take on V(D)J recombination: transcription driven nuclear and chromatin reorganization in RAG-mediated cleavage</i></b>	Julie Chaumeil and Jane A. Skok
<b>16</b>	<b><i>CTCF and ncRNA regulate the three-dimensional structure of antigen receptor loci to facilitate V(D)J recombination</i></b>	Nancy M. Choi and Ann J. Feeney
<b>24</b>	<b><i>Ubiquitination events that regulate recombination of immunoglobulin loci gene segments</i></b>	Jaime Chao, Gerson Rothschild and Uttiya Basu
<b>36</b>	<b><i>Function of YY1 in long-distance DNA interactions</i></b>	Michael L. Atchison
<b>47</b>	<b><i>Balancing proliferation with Ig<math>\kappa</math> recombination during B-lymphopoiesis</i></b>	Keith M. Hamel, Malay Mandal, Sophiya Karki and Marcus R. Clark
<b>56</b>	<b><i>A novel Pax5-binding regulatory element in the Ig<math>\kappa</math> locus</i></b>	Rena Levin-Klein, Andrei Kirillov, Chaggai Rosenbluh, Howard Cedar and Yehudit Bergman
<b>67</b>	<b><i>The interplay between chromatin and transcription factor networks during B cell development: who pulls the trigger first?</i></b>	Mohamed Amin Choukrallah and Patrick Matthias
<b>78</b>	<b><i>The Bright side of hematopoiesis: regulatory roles of ARID3a/Bright in human and mouse hematopoiesis</i></b>	Michelle L. Ratliff, Troy D. Templeton, Julie M. Ward and Carol F. Webb
<b>86</b>	<b><i>Architecture and expression of the Nfatc1 gene in lymphocytes</i></b>	Ronald Rudolf, Rhoda Busch, Amiya K. Patra, Khalid Muhammad, Andris Avots, Jean-Christophe Andrau, Stefan Klein-Hessling and Edgar Serfling
<b>90</b>	<b><i>Genomic architecture may influence recurrent chromosomal translocation frequency in the Igh locus</i></b>	Amy L. Kenter, Robert Wuerffel, Satyendra Kumar and Fernando Grigera
<b>95</b>	<b><i>AIDing chromatin and transcription-coupled orchestration of immunoglobulin class-switch recombination</i></b>	Bharat Vaidyanathan, Wei-Feng Yen, Joseph N. Pucella and Jayanta Chaudhuri
<b>108</b>	<b><i>Epigenetic regulation of individual modules of the immunoglobulin heavy chain locus 3' regulatory region</i></b>	Barbara K. Birshstein



**117 *Oct2 and Obf1 as facilitators of B:T cell collaboration during a humoral immune response***

Lynn Corcoran, Dianne Emslie, Tobias Kratina, Wei Shi, Susanne Hirsch, Nadine Taubenheim and Stephane Chevrier

**130 *Regulation of the NF- $\kappa$ B-mediated transcription of inflammatory genes***

Dev Bhatt and Sankar Ghosh

**139 *CIITA and its dual roles in MHC gene transcription***

Ballachanda N. Devaiah and Dinah S. Singer



# Chromatin and transcriptional tango on the immune dance floor

Ananda L. Roy<sup>1\*</sup> and Robert G. Roeder<sup>2\*</sup>

<sup>1</sup> Programs in Immunology and Genetics, Department of Developmental, Molecular and Chemical Biology, Tufts University School of Medicine, Boston, MA, USA

<sup>2</sup> Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY, USA

\*Correspondence: ananda.roy@tufts.edu; roeder@mail.rockefeller.edu

**Edited and reviewed by:**

Thomas L. Rothstein, The Feinstein Institute for Medical Research, USA

**Keywords: immune response, transcription, promoter, enhancer, chromatin**

The process of generating differentiated cell types performing specific effector functions from their respective undifferentiated precursors is dictated by extracellular signals, which alter the host cell's capacity to perform cellular functions. One major mechanism for bringing about such changes is at the level of transcription. Thus, the transcription-related induction of previously silent genes and suppression of active genes in response to extracellular signals can result in the acquisition of new functions by the cells. The general transcriptional machinery, which comprised of RNA Polymerase II and associated initiation factors, assemble into preinitiation complexes at the core promoters of eukaryotic protein coding genes in response to the signal-dependent activation of corresponding regulatory factors that bind to promoter and enhancer elements (1). The rate of formation and/or stability of these complexes, which can be modulated both by enhancer–promoter interactions and by chromatin structural modifications, dictate the transcriptional regulation of the corresponding gene. Such coordinated temporal and spatial regulation of gene expression in response to specific signals determines lineage differentiation, cellular proliferation, and development (2).

Every event in the life cycle of a lymphocyte is modulated by the signals they receive. For instance, expression of the B cell antigen receptor (BCR) on the surface of B cells is a hallmark of various stages of B cell development, with signaling through the BCR being important during both early/antigen-independent (tonic) and late/antigen-dependent phases of development (3). However, how BCR signaling connects to chromatin changes and downstream transcriptional pathways at each step of development remains poorly understood. Similar questions also remain in other cells of the immune system. In particular, how enhancers communicate with promoters in a stage-specific fashion and in the context of chromatin also remain unclear (2). Chromatin modifiers are generally present and active in most cell types (4, 5). How then could there be gene-specific differences in chromatin architecture dependent on a particular stage of development?

The B (and T) lymphocytes also perform a unique developmental program because they have an unparalleled genetic makeup – the genetic loci that encode their cell surface receptors are in an “unrearranged” or “germline” configuration during the early stages of development. Thus, while expressing stage-specific genes and transcription factors during each developmental stage, lymphocytes also need to undergo rearrangement of their cognate receptor loci in a strictly ordered fashion to generate a pool of

receptor proteins that, individually, are capable of recognizing specific antigens that are encountered at a much later step (6). Hence, there must be a strict negotiation between the recombination machinery and the transcriptional machinery at every developmental step. Importantly, along the way, those B cells that express receptors capable of recognizing self-antigens must be eliminated to avoid autoimmune responses and only those cells capable of recognizing foreign-antigens are preserved for migration to peripheral organs where they eventually encounter pathogens. How are these processes coordinately regulated in a stage-specific fashion and what role does chromatin play? Are the rules of engagement different in innate versus adaptive immune responses? The following 15 articles address some of these questions and provide important insights regarding our current understanding of signal-induced chromatin and transcriptional regulation of the immune system.

## REGULATION OF V(D)J RECOMBINATION – ROLE OF TRANSCRIPTION AND CHROMATIN

Germline configurations of antigen receptor loci in B and T lymphocytes have hundreds of variable (V) region gene-segments, which have the potential to combine with a select few diversity (D) and joining (J) gene-segments to create recombined genes encoding numerous receptors that can recognize a vast repertoire of antigens (6, 7). Given the importance and timing of these events, it is no wonder that the process of “V(D)J recombination” is exquisitely regulated at multiples levels. Two exciting articles, one by Chaumeli and Skok (8) and the other by Choi and Feeney (9), review our current understanding of how transcription factors, chromatin architecture, and the three-dimensional architecture of the nucleus and the topology of genomic DNA regulate this process. An interesting article by Basu and colleagues describes how ubiquitination events regulate the RAG and activation-induced cytidine deaminase (AID) enzymes that are important for recombination (10). Moreover, this article also discusses how these post-translational events also regulate DNA damage at undesirable loci and during cell cycle phases (10).

## TRANSCRIPTION FACTORS IN HEMATOPOIETIC DEVELOPMENT

Recombination and transcription are coupled during hematopoietic development (11–13). The next set of articles deal with factors involved in this coordination. Atchison and colleagues describes

the role of an important but ubiquitously expressed transcription factor YY1 in this highly tissue-specific function (14). Clark and colleagues review the function of interleukin-7 receptor (IL7R) and transcription factor STAT5 in balancing proliferation and recombination of the immunoglobulin light chain (Ig $\kappa$ ) gene (15). Bergman and colleagues present primary studies on the role of another essential transcription factor Pax5 in regulating the Ig $\kappa$  gene (16). The sequential involvement of transcription factors and chromatin regulators remains an open question, and Choukrallah and Matthias review our current understanding of these factors in B cell development (17). Webb and colleagues discuss the role of transcription factor Bright in both human and mouse B cell development (18), while Serfling and colleagues review the role of NFATc1 transcription factor during hematopoiesis (19).

### REGULATION OF CLASS-SWITCH RECOMBINATION AND SOMATIC HYPERMUTATION

Because mature B cells encounter a variety of antigens, they undergo both Class-Switch recombination (CSR) and somatic hypermutation (SHM) to diversify their antibody repertoire by utilizing enzymes such as AID. Given that these processes involve DNA breaks, they must be extremely tightly regulated to maintain genomic integrity (20, 21). Kenter and colleagues (22) and Chaudhuri and colleagues (23) present two articles discussing various factors regulating both SHM and CSR, including three-dimensional genomic topology, chromatin, and transcription. Barbara Birshtein discusses the role of the 3'-enhancer in controlling both SHM and CSR, in particular the epigenetic architecture of the enhancer in these processes (24).

### TRANSCRIPTION FACTORS REGULATING IMMUNE RESPONSES

The ultimate role of immune cells is to mount an effective adaptive or innate response against pathogens (25). Hence, the transcription factors regulating these responses play an extremely important role. The final three articles deal with the transcription factors involved in immune responses and antigen presentation. Corcoran and colleagues present primary data on the function of transcription factor Oct2 and its co-activator Obf1/OCA-B in collaboration between B and T cells during an adaptive immune response (26). Bhatt and Ghosh discuss the role of the critical transcription factor NF- $\kappa$ B in innate immune response and how it controls the process of inflammation, which is crucial in maintaining immune homeostasis (27). Finally, Devaiah and Singer discuss our current understanding of the role of Class II transactivator CIITA (28), which is a master regulator of major histocompatibility complex gene expression necessary for antigen presentation (29).

### PERSPECTIVE

Mechanisms that regulate communication between enhancers and promoters are complex and involve many transcription factors, accessory molecules and chromatin regulators (30). Given the exquisite timing and precision that are necessary to mount an effective immune response, it is fully anticipated that such complex regulatory mechanisms must be in full display for this to occur. The next few years will undoubtedly uncover more surprises

that ultimately will lead to a better understanding of the role of transcription in immune responses.

### REFERENCES

1. Roeder RG. Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett* (2005) **579**:909–15. doi:10.1016/j.febslet.2004.12.007
2. Roy AL, Sen R, Roeder RG. Enhancer-promoter communication and transcriptional regulation of Igh. *Trends Immunol* (2011) **32**(11):532–9. doi:10.1016/j.it.2011.06.012
3. Kurosaki T, Shinohara H, Baba Y. B cell signaling and fate decision. *Annu Rev Immunol* (2010) **28**:21–55. doi:10.1146/annurev.immunol.021908.132541
4. Guertin MJ, Lis JT. Mechanisms by which transcription factors gain access to target sequence elements in chromatin. *Curr Opin Genet Dev* (2013) **23**(2):116–23. doi:10.1016/j.gde.2012.11.008
5. Fowler T, Sen R, Roy AL. Regulation of primary response genes. *Mol Cell* (2011) **44**(3):348–60. doi:10.1016/j.molcel.2011.09.014
6. Bossen C, Mansson R, Murre C. Chromatin topology and the regulation of antigen receptor assembly. *Annu Rev Immunol* (2012) **30**:337–56. doi:10.1146/annurev-immunol-020711-075003
7. Sleckman BP, Oltz EM. Preparing targets for V(D)J recombinase: transcription paves the way. *J Immunol* (2012) **188**(1):7–9. doi:10.4049/jimmunol.1103195
8. Chaumeil J, Skok JA. A new take on v(d)j recombination: transcription driven nuclear and chromatin reorganization in rag-mediated cleavage. *Front Immunol* (2013) **4**:423. doi:10.3389/fimmu.2013.00423
9. Choi NM, Feeney AJ. CTCF and ncRNA regulate the three-dimensional structure of antigen receptor loci to facilitate V(D)J recombination. *Front Immunol* (2014) **5**:49. doi:10.3389/fimmu.2014.00049
10. Chao J, Rothschild G, Basu U. Ubiquitination events that regulate recombination of immunoglobulin Loci gene segments. *Front Immunol* (2014) **5**:100. doi:10.3389/fimmu.2014.00100
11. Johnson K, Reddy KL, Singh H. Molecular pathways and mechanisms regulating the recombination of immunoglobulin genes during B-lymphocyte development. *Adv Exp Med Biol* (2009) **650**:133–47. doi:10.1007/978-1-4419-0296-2\_11
12. Santos PM, Borghesi L. Molecular resolution of the B cell landscape. *Curr Opin Immunol* (2011) **23**(2):163–70. doi:10.1016/j.coi.2010.11.014
13. Naito T, Tanaka H, Naoe Y, Taniuchi I. Transcriptional control of T-cell development. *Int Immunol* (2011) **23**(11):661–8. doi:10.1093/intimm/dxr078
14. Atchison ML. Function of YY1 in long-distance DNA interactions. *Front Immunol* (2014) **5**:45. doi:10.3389/fimmu.2014.00045
15. Hamel KM, Mandal M, Karki S, Clark MR. Balancing proliferation with Ig $\kappa$  recombination during B-lymphopoiesis. *Front Immunol* (2014) **5**:139. doi:10.3389/fimmu.2014.00139
16. Levin-Klein R, Kirillov A, Rosenbluh C, Cedar H, Bergman Y. A novel pax5-binding regulatory element in the ig $\kappa$  locus. *Front Immunol* (2014) **5**:240. doi:10.3389/fimmu.2014.00240
17. Choukrallah MA, Matthias P. The interplay between chromatin and transcription factor networks during B cell development: who pulls the trigger first? *Front Immunol* (2014) **5**:156. doi:10.3389/fimmu.2014.00156
18. Ratliff ML, Templeton TD, Ward JM, Webb CF. The bright side of hematopoiesis: regulatory roles of ARID3a/bright in human and mouse hematopoiesis. *Front Immunol* (2014) **5**:113. doi:10.3389/fimmu.2014.00113
19. Rudolf R, Busch R, Patra AK, Muhammad K, Avots A, Andrau JC, et al. Architecture and expression of the nfatc1 gene in lymphocytes. *Front Immunol* (2014) **5**:21. doi:10.3389/fimmu.2014.00021
20. Li G, Zan H, Xu Z, Casali P. Epigenetics of the antibody response. *Trends Immunol* (2013) **34**(9):460–70. doi:10.1016/j.it.2013.03.006
21. Kato L, Stanlie A, Begum NA, Kobayashi M, Aida M, Honjo T. An evolutionary view of the mechanism for immune and genome diversity. *J Immunol* (2012) **188**(8):3559–66. doi:10.4049/jimmunol.1102397
22. Kenter AL, Wuerffel R, Kumar S, Grigera F. Genomic architecture may influence recurrent chromosomal translocation frequency in the Igh locus. *Front Immunol* (2013) **4**:500. doi:10.3389/fimmu.2013.00500
23. Vaidyanathan B, Yen WF, Pucella JN, Chaudhuri J. AIDing chromatin and transcription-coupled orchestration of immunoglobulin class-switch recombination. *Front Immunol* (2014) **5**:120. doi:10.3389/fimmu.2014.00120
24. Birshtein BK. Epigenetic regulation of individual modules of the immunoglobulin heavy chain locus 3' regulatory region. *Front Immunol* (2014) **5**:163. doi:10.3389/fimmu.2014.00163

25. Smale ST, Tarakhovsky A, Natoli G. Chromatin contributions to the regulation of innate immunity. *Annu Rev Immunol* (2014) **32**:489–511. doi:10.1146/annurev-immunol-031210-101303
26. Corcoran L, Emslie D, Kratina T, Shi W, Hirsch S, Taubenheim N, et al. Oct2 and Obf1 as facilitators of B:T cell collaboration during a humoral immune response. *Front Immunol* (2014) **5**:108. doi:10.3389/fimmu.2014.00108
27. Bhatt D, Ghosh S. Regulation of the NF- $\kappa$ B-mediated transcription of inflammatory genes. *Front Immunol* (2014) **5**:71. doi:10.3389/fimmu.2014.00071
28. Devaiah BN, Singer DS. CIITA and its dual roles in MHC gene transcription. *Front Immunol* (2013) **4**:476. doi:10.3389/fimmu.2013.00476
29. Harton JA, Ting JP. Class II transactivator: mastering the art of major histocompatibility complex expression. *Mol Cell Biol* (2000) **20**(17):6185–94. doi:10.1128/MCB.20.17.6185-6194.2000
30. Zhang Y, Wong CH, Birnbaum RY, Li G, Favaro R, Ngan CY, et al. Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. *Nature* (2013) **504**(7479):306–10. doi:10.1038/nature12716

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 October 2014; accepted: 25 November 2014; published online: 15 December 2014.

Citation: Roy AL and Roeder RG (2014) Chromatin and transcriptional tango on the immune dance floor. *Front. Immunol.* **5**:631. doi: 10.3389/fimmu.2014.00631

This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Roy and Roeder. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# A new take on V(D)J recombination: transcription driven nuclear and chromatin reorganization in RAG-mediated cleavage

Julie Chaumeil<sup>†</sup> and Jane A. Skok<sup>\*</sup>

Department of Pathology, New York University School of Medicine, New York, NY, USA

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

Claude-Agnes Reynaud, INSERM, France

Rodney P. DeKoter, The University of Western Ontario, Canada

## \*Correspondence:

Jane A. Skok, Department of Pathology, New York University School of Medicine, 550 First Avenue, MSB 599, New York, NY 10016, USA  
e-mail: jane.skok@nyumc.org

## <sup>†</sup>Present address:

Julie Chaumeil, Mammalian Developmental Epigenetics Group, Institut Curie, CNRS UMR 3215, INSERM U934, Paris, France

It is nearly 30 years since the Alt lab first put forward the accessibility model, which proposes that cleavage of the various antigen receptor loci is controlled by lineage and stage specific factors that regulate RAG access. Numerous labs have since demonstrated that locus opening is regulated at multiple levels that include sterile transcription, changes in chromatin packaging, and alterations in locus conformation. Here we focus on the interplay between transcription and RAG binding in facilitating targeted cleavage. We discuss the results of recent studies that implicate transcription in regulating nuclear organization and altering the composition of resident nucleosomes to promote regional access to the recombinase machinery. Additionally we include new data that provide insight into the role of the RAG proteins in defining nuclear organization in recombining T cells.

**Keywords:** V(D)J recombination, transcription, nuclear organization, higher-order loops, ATM, nucleosomes, RAG, pericentromeric heterochromatin

## INTRODUCTION

V(D)J recombination occurs during lymphocyte development to create B and T cell receptors that can recognize a vast array of foreign antigen. Diversity is generated within the seven antigen-receptor loci (four T cell receptor loci, *Tcrg*, *Tcrd*, *Tcrb*, and *Tcra* and three immunoglobulin loci, *Igh*, *Igk*, and *Igl*) by reshuffling variable (V), diversity (D), and joining (J) gene segments that are arrayed along the length of each of these large loci. Rearrangement is mediated by the RAG recombinase, which binds to highly conserved heptamer and nonamer recombination signal sequences (RSSs) that flank each of the V, D, and J gene segments. Both RAG1 and RAG2 proteins, which make up the recombinase, bind to two segments, bringing them together to form a synapse prior to the introduction of double strand breaks (DSBs). In addition, RAG plays a role after cleavage by holding the four broken ends together in a RAG post cleavage complex that directs repair by the ubiquitous classical non-homologous end joining (C-NHEJ) pathway.

Although the process of rearrangement is common to all antigen-receptor loci and mediated by the same machinery, it is regulated so that *Ig* and *Tcr* loci are respectively rearranged at the appropriate stage of B and T cell development. Furthermore, cleavage is restricted at the allelic level (allelic exclusion) to ensure rearrangement and cell surface expression of a single specificity receptor. Studies from numerous labs have validated the accessibility model put forward by the Alt lab and shown that rearrangement is linked with transcription, active histone modifications, and reversible locus contraction, which brings widely separated gene segments together by looping (1, 2).

## RAG BINDING AND DISTRIBUTION WITHIN THE NUCLEUS

Both RAG1 and RAG2 are required for cleavage although the endolytic activity lies within the RAG1 protein. RAG2 binds to chromatin via its PHD domain, which specifically recognizes the histone modification, H3K4me3 (3, 4). Genome wide ChIP-seq analyses indicate that RAG2 recruitment mirrors the footprint of this active histone modification. In contrast, RAG1 binding is more directed and occurs predominantly at conserved RSSs (5), however binding can also occur at cryptic RSS sites that are scattered throughout the genome. As RAG binding is not limited to the antigen-receptor loci alone, this raises a question about the mechanisms that direct cleavage. Clearly, DSBs are not introduced everywhere in the genome at sites of active chromatin or indeed at consensus/cryptic RSSs, so there must be other factors involved in determining when breaks are generated.

One possibility to consider, beyond active chromatin and the nature of the RSS, is the localized concentration of RAG1/2. It is logical to assume that the higher the concentration of recombinase in the vicinity of a vulnerable gene, the more likely the chances of cleavage. RAG2 localizes to euchromatic regions of the nucleus and domains of RAG enrichment are clearly visible by microscopy after immunostaining (Hewitt and Skok, unpublished). But what is the mechanism underlying the generation of these focal centers? The data from recent genome wide chromosome conformation capture experiments indicates that co-regulated actively transcribed genes come together in the nucleus in transcription factories (6, 7). Thus, contact between common transcription factor or RAG bound loci will likely increase the local concentration of these factors in the nucleus, as shown for polycomb bound regions that associate to

form a polycomb body (8). Since gene expression depends on the integrated binding of a number of different remodeling and transcription factors, the balance of these will likely determine which factors are dominant in defining the intra- and inter-chromosomal interaction partners of any particular locus.

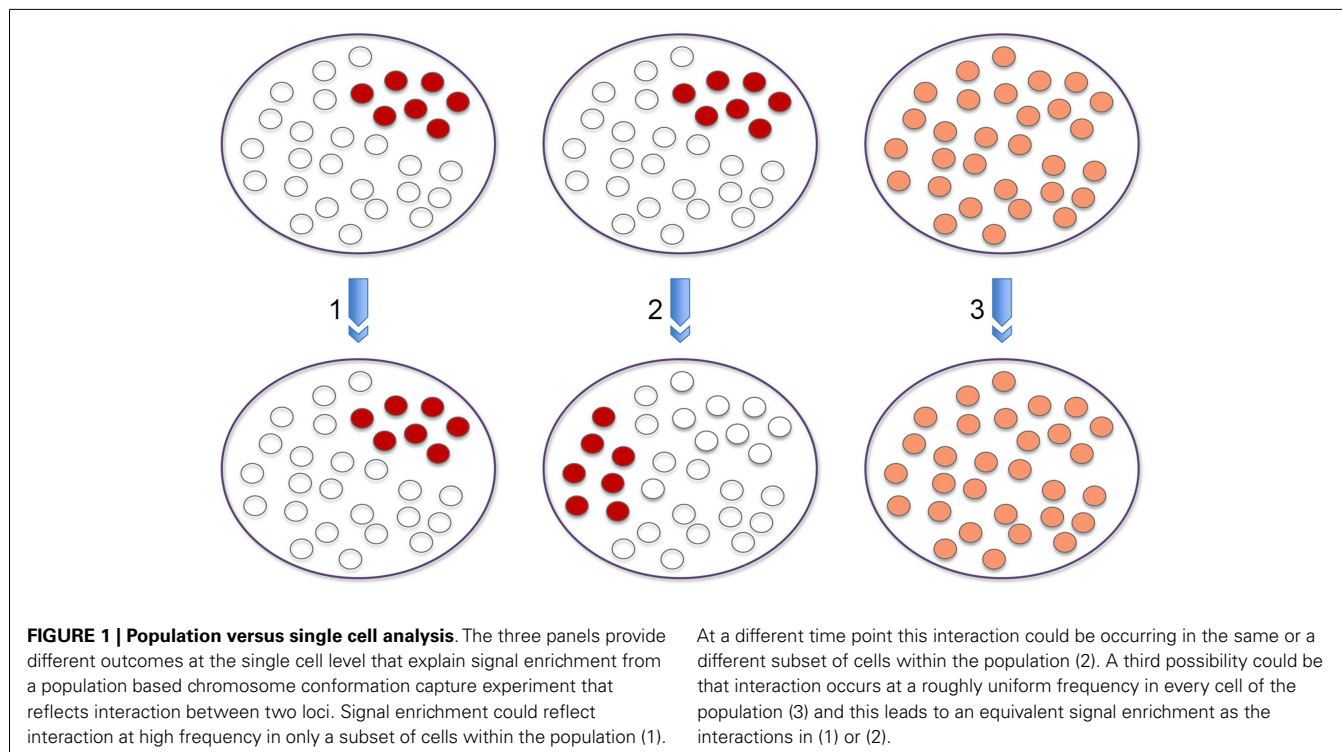
### POPULATION VERSUS SINGLE CELL ANALYSIS

When considering the data from genome wide association studies it is important to remember that signal enrichment reflects the sum of the data derived from a population of cells. What happens at the single cell level may be very different. However, without live systems in which we can track the movements of individual loci in single cells over a period of time, at the simplest level, when focusing on an interaction between two loci in a population, we cannot tell whether chromosome conformation capture signal enrichment reflects interaction at high frequency in only a subset of cells within the population (1); whether at a different time point this interaction will be occurring in the same (1) or a different subset of cells within the population (2); or whether interaction occurs at a roughly uniform frequency in every cell of the population (3) and this leads to an equivalent signal enrichment as the interactions in (1) or (2) (**Figure 1**). Nonetheless, in the absence of these live systems we can address some of these questions using single cell DNA FISH analyses on a population of fixed cells. With this approach, although we cannot distinguish between (1) and (2), we can distinguish between these two alternatives and (3) by determining whether interactions occur at a similar frequency in every cell versus a high frequency in a subset of cells. The same issues arise for histone modifications and transcription factor/RAG binding at a particular site. Furthermore, if genome wide data sets are integrated, e.g., chromosome

conformation capture and ChIP-seq, the situation becomes even more complex.

### THE ROLE OF RAG IN INTER-CHROMOSOMAL INTERACTIONS

To examine these issues in the context of V(D)J recombination, we asked whether RAG binding could have a role in bringing RAG bound antigen-receptor loci together in the nucleus in localized recombination centers. We discovered that expression of RAG1 brings target homologous antigen-receptor alleles together in a subset of recombining cells (9–11). Homologous pairing of *Ig* or *Tcr* alleles occurs prior to and independent of RAG cleavage because expression of a catalytically inactive RAG1D708A mutant protein can rescue pairing in RAG1-deficient cells. In addition to increasing the local concentration of RAG in the nucleus, a second not mutually exclusive possibility is that communication between the two alleles could be important for regulation of cleavage on homologs. Indeed, we found that the introduction of a break on one allele halts the introduction of further breaks on the second allele through the action of the DNA damage response factor Ataxia telangiectasia mutated (ATM) (10, 11). Briefly, ATM, recruited to the site of a break on one allele, acts *in trans* on the second allele repositioning it to pericentromeric heterochromatin (PCH). Transient relocation to this repressive nuclear environment likely causes a degree of silencing that depletes RAG binding on the uncleaved allele during repair of the first break. Thus, ATM-mediated changes in nuclear organization function to ensure asynchronous RAG-mediated cleavage on homologous alleles. Regulation *in trans* is important for the initiation of allelic exclusion and for restricting the number of DSBs that are introduced at any one time in the cell (10). Based on our results we favor a model in which RAG-mediated breaks are introduced on closely



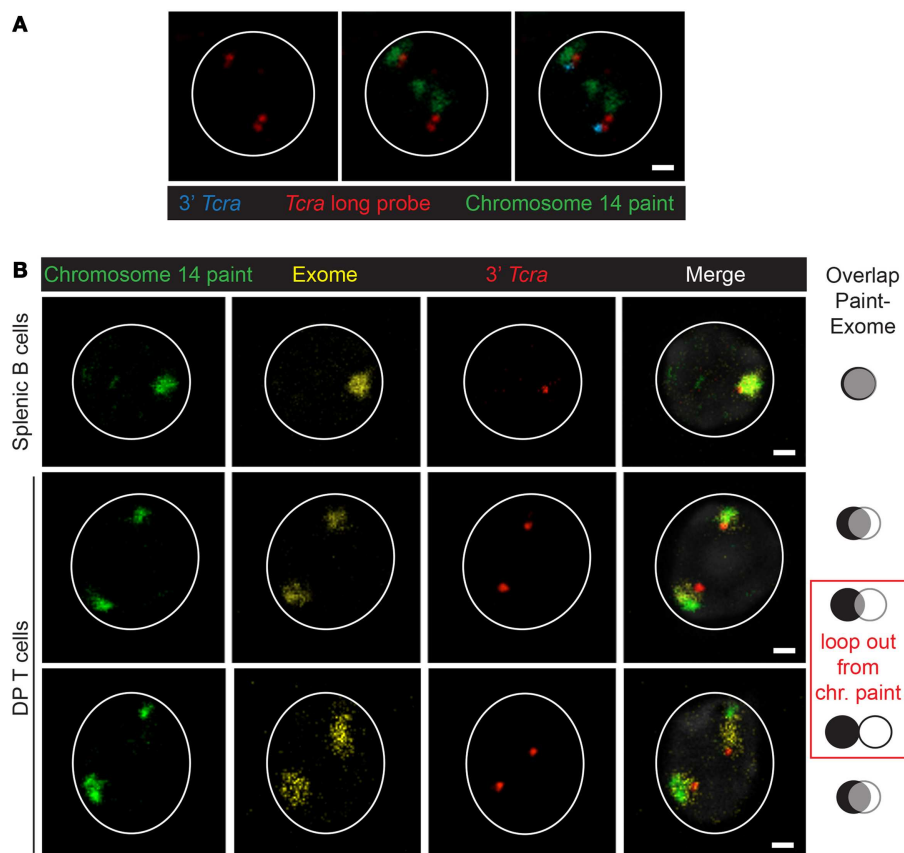


associated homologs and then separate for repair to facilitate regulated asynchronous cleavage. However, without a live imaging system in which we can track the dynamics of cleavage and repair, we cannot definitively determine whether this is the case. Nevertheless, it is clear that if homologs are paired the uncleaved allele will have immediate access to a high concentration of activated ATM recruited to the site of damage on the cleaved allele.

### HIGHER-ORDER LOOP FORMATION DURING RECOMBINATION

Beyond locus contraction and homolog pairing we recently uncovered an additional layer of regulation involving nuclear organization that occurs during V(D)J recombination: the formation of higher-order loops (11). Chromosomes occupy discreet territories in interphase cells and the size and position of these within the nucleus is dependent on the cell type and developmental stage (12). Live imaging studies have shown that chromosome territories move very little following mitosis (13), so gene mobility facilitated by the formation of higher-order loops provides an opportunity for loci on different chromosomes to contact each other in nuclear space. Movement of genes away from their individual

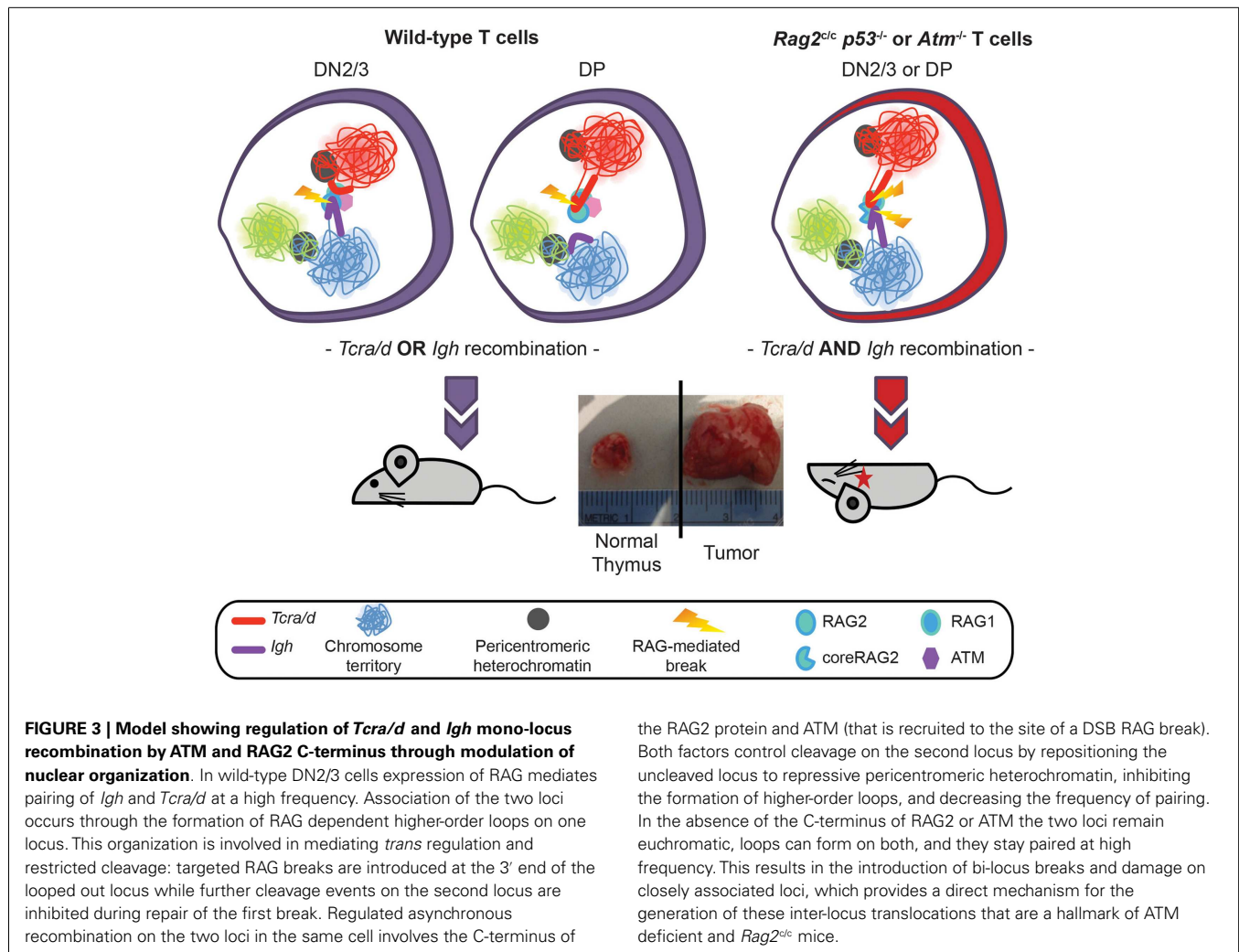
chromosome territories linked to activation/transcription (14) has been shown to facilitate stochastic inter-chromosomal interactions (15), but little is known about whether pairing of this sort could be involved in regulation of genes *in trans*. Our data indicate that, as with homolog pairing, RAG1 expression (independent of its catalytic activity) induces the formation of higher-order loops that separate the 3' end of the antigen-receptor locus, *Tcra*, from its 5' end which remains embedded in the chromosome territory (as assessed by DNA FISH with a chromosome paint probe) (11) (Figure 2A). Furthermore, *Tcra* expression is linked to looping and pairing because in splenic B cells where RAG is not present and *Tcra* is not transcribed loop formation is inhibited and the two loci pair at a frequency below the levels seen in RAG1-deficient cells. Additional RNA/DNA FISH analyses revealed that the proportion of looped out alleles that are transcribed is greater than those located at the outer edge of the chromosome territory, while those alleles that are buried in the territory are not associated with any RNA signal at all. It should be noted that although we used an oligonucleotide probe pool covering the entire locus except for the most repetitive regions, nascent RNA signals could only be detected at the 3' end of *Tcra*, likely because with this assay



**FIGURE 2 | Organization of the chromosome 14. (A)** Example of looping out of *Tcra/d* from the chromosome 14 territory in wild-type DP cells. Chromosome 14 paint in green, long *Tcra/d* probe in red, 3' *Tcra* probe (BAC RP23-255N13) in blue. Scale bar = 1  $\mu$ m.

**(B)** Examples showing the 3D-organization of the chromosome 14 territory with a chromosome paint (repetitive sequences) in green, an

exome probe (gene-rich regions) in yellow and the 3' *Tcra* probe in red. The paint and exome domains overlap in splenic B cells and in DP T cells when *Tcra* is close to the paint domain and not looped out, while the two signals are more separated when *Tcra* is looped out from the paint. In either case *Tcra* is never looped out from the exome domain. Scale bar = 1  $\mu$ m.



there is a threshold below which transcription cannot be detected. Furthermore, we found that RNA signals are not distributed uniformly throughout the population: 1/3 of the cells had no *Tcr* signals, 1/3 of the cells had one *Tcr* signal and the remaining third had signals from both alleles. Thus, *Tcr* transcription of 3' *Tcr* occurs in a subset of cells and this is linked with looping out of this region.

To examine in further detail how the chromosome territory and the locus are organized in nuclear space, we designed an oligonucleotide probe pool covering only the exon sequences represented on chromosome 14 (called "exome") (16). We used the exome in conjunction with a conventional chromosome 14 paint (that largely encompasses repetitive sequences), as well as the 3' *Tcr* BAC probe. Curiously, we found considerable overlap between the paint and exome DNA probes for chromosome 14 in splenic B cells and in DP T cells on chromosomes where *Tcr* is not looped out (Figure 2B). In contrast there was very little overlap between the paint and exome in DP cells on the allele on which *Tcr* forms higher-order loops (Figure 2B). These observations indicate that looping out of *Tcr* extends exome sequences, dragging them to the outside of the territory so that they no longer overlap with the

paint signal. These data, underline the link between loop formation and transcription that we and others previously documented (11, 14, 17). The Bickmore lab previously showed that chromosome paints only reveal the core of chromosome territories while exome probes can detect gene-rich regions that are mostly located around the outside of these core domains (16). Here our analysis of chromosome 14 in two different cell types provides additional insights into the dynamics of chromosome organization, indicating that high levels of expression from an individual locus and the presence of a *trans* acting factor such as RAG, can impact on looping, with the gene-rich regions being reorganized outside of the core domain alongside the highly transcribed *Tcr*.

What is the purpose of higher-order loop formation in DP T cells? Interestingly, we found that looping out of *Tcr* occurs on only one allele and this was linked to the occurrence of RAG-mediated mono-allelic breaks on the 3' end of the looped out of *Tcr* (11). These data link up with previous genome wide analyses from the Schatz lab showing that active histone marks and RAG1/2 binding are enriched at the 3' region of recombining antigen-receptor loci (5). However, these are population derived data so they do not distinguish between signals in individual cells and on



homologous alleles, thus it is not possible to determine whether RAG and active histone marks are enriched on antigen-receptor loci in a subset of cells and specifically whether RAG binding is more concentrated on the looped out allele. Nonetheless, it is clear that RAG binding coupled with a high level of transcription is linked to movement of the 3' end of one allele away from the territory and this in turn correlates with the introduction of RAG-mediated mono-allelic cleavage on the looped out allele.

### REGULATION OF MONO-LOCUS CLEAVAGE

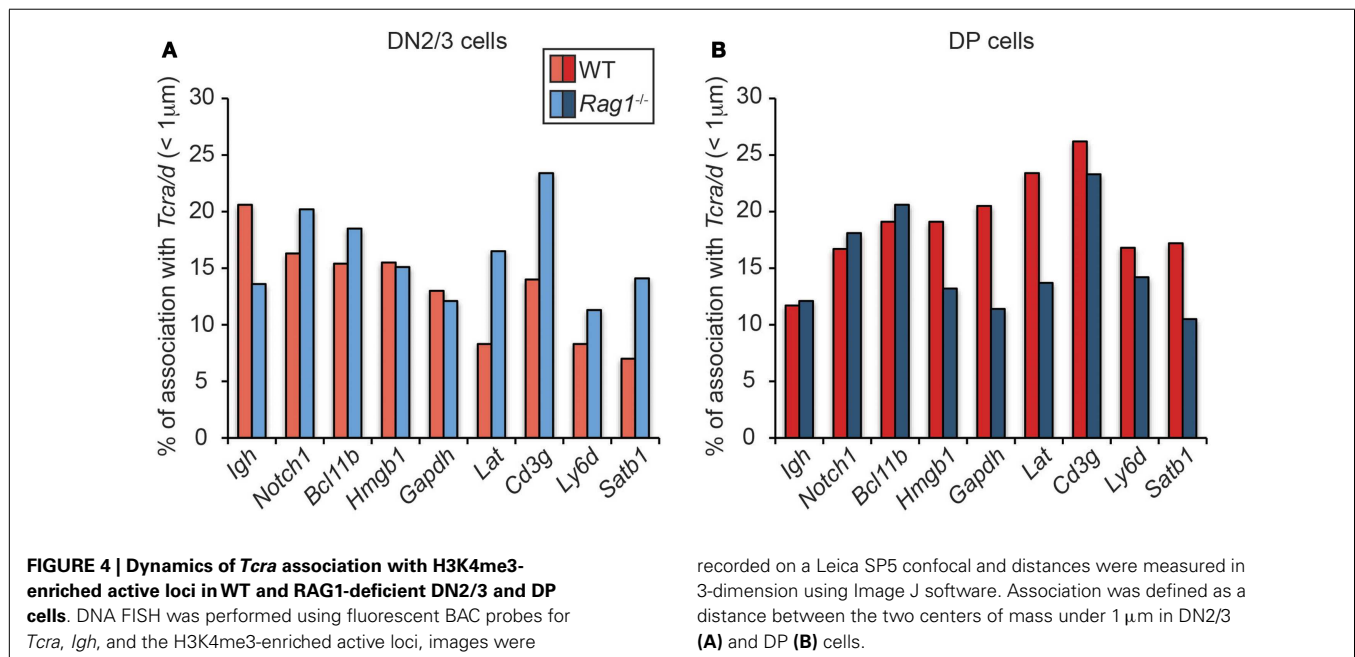
To extend these studies we asked whether similar mechanisms could control RAG cleavage on different loci undergoing recombination at the same stage in development. For this we analyzed the *Tcra/d* locus in conjunction with the *Igh* locus. Although *Tcrd* and *Tcra* occupy the same chromosomal location (*Tcrd* is embedded within the *Tcra* locus) the two loci undergo recombination at different stages of T cell development in CD4 CD8 double negative, DN2/3, and CD4 CD8 double positive DP cells, respectively. For reasons that are not well understood, the *Igh* locus undergoes a low level of partial rearrangement in T lineage cells and in this context it is of note that *Igh* has been identified as a translocation partner of *Tcra/d* in T lineage lymphomas. Specifically, we and others have found that *Tcra/d-Igh* translocations occur in T-lymphomas from ATM deficient mice (18–21). Furthermore, we recently discovered T-lymphomas with *Tcra/d-Igh* translocations in mice expressing a truncated version of RAG2, missing the non-core regulatory C-terminal domain crossed onto a p53 deficient background (18). Our recent investigations have revealed that ATM and the RAG2 C-terminus prevent bi-locus RAG-mediated cleavage through similar mechanisms: modulation of three-dimensional conformation (higher-order loops) and nuclear organization of the two loci (22). Thus, the RAG2 C-terminus and ATM control asynchronous RAG cleavage on homologous and heterologous antigen-receptor alleles in a similar manner: through repositioning the uncleaved allele/locus at repressive PCH, inhibiting

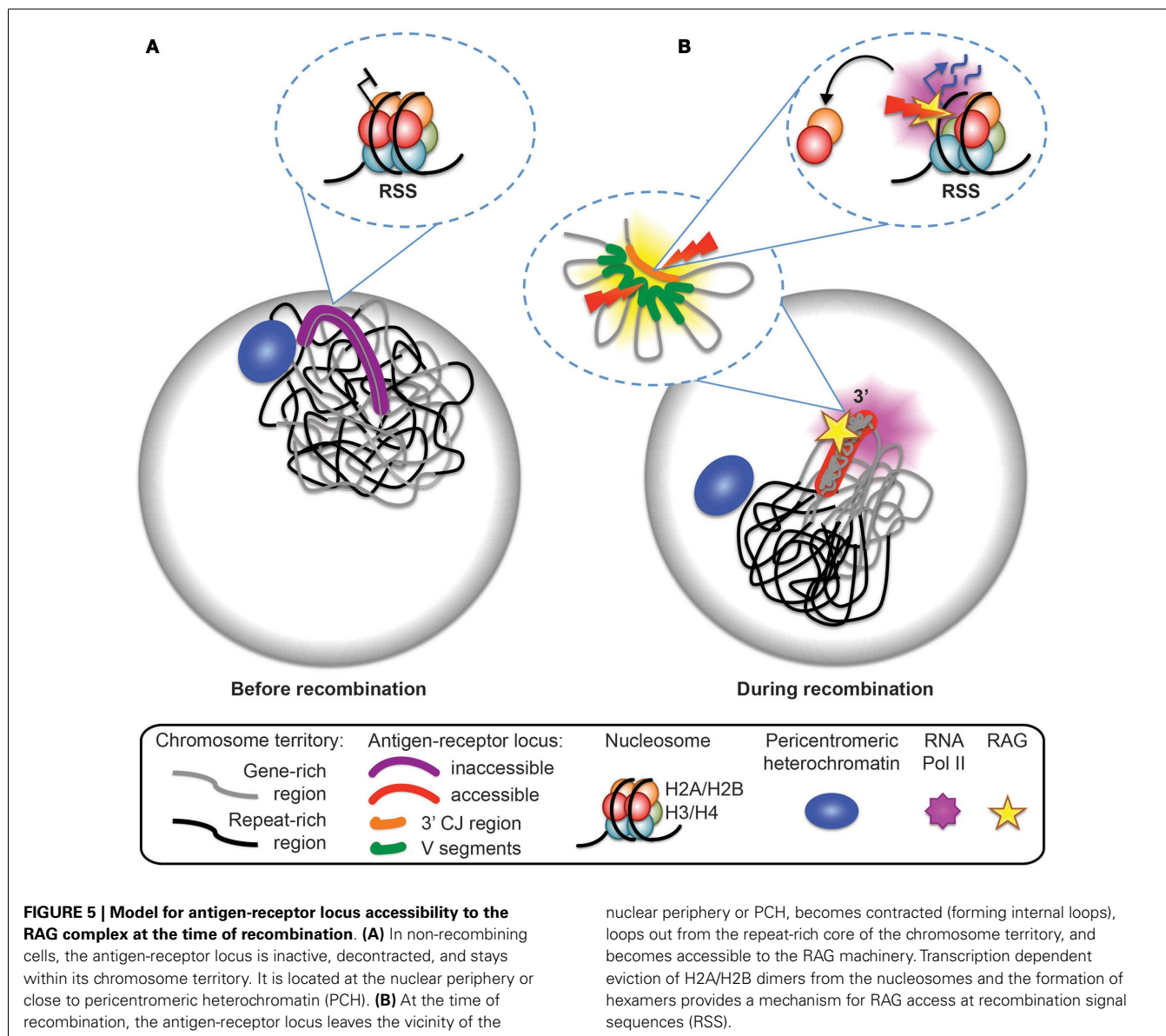
bi-allelic/bi-locus loop formation and bi-allelic/bi-locus cleavage. This limits the number of potential substrates for translocation and provides an important mechanism for protecting genome stability (Figure 3) (11, 22).

### RAG BRINGS RECOMBINING LOCI TOGETHER IN THE NUCLEUS

Interestingly, control of cleavage of *Tcra/d* and *Igh* in T cells involves RAG-mediated regulation of association of the two loci. Thus, it appears that RAG brings recombining loci together in the nucleus. In this case the two loci are close together in DN cells (when both *Tcrd* and *Igh* are recombining) but interact far less frequently in DP cells (when *Igh* is recombined at lower levels). This raises the possibility that feedback control of RAG activity could involve close communication between heterologous antigen-receptor alleles as we have proposed for homologous alleles. Either way, it appears that regulation *in trans* involves control of loop formation and repositioning of the other allele/locus to a repressive compartment of the nucleus.

Given these findings and the fact that RAG2 can bind to H3K4me3-enriched active loci throughout the genome, we considered the possibility that there could be localized feedback control through association of RAG enriched loci. However, when we looked to see whether RAG induces association of *Tcra/d* with other hematopoietic lineage specific or housekeeping H3K4me3-enriched loci on different chromosomes in DN2/3 cells (where *Tcrd* and *Igh* are recombining), we found an opposite trend to the relationship between *Igh* and *Tcra/d* (Figure 4A), and as shown previously (22): depletion of RAG1 in the cells increased rather than decreased association of these other genes with *Tcra/d*. This trend was mirrored by *Notch1* and *Bcl11b* that we newly examined here: if anything, RAG1 expression separates these loci from recombining *Tcrd* at the DN2/3 stage of development. This is of interest because RAG targeting of both these loci is linked to lymphoid malignancies (23). In contrast, when we examined





interactions in DP cells (where *Tcra* is recombining) while RAG depletion did not alter the relationship between *Igh* and *Tcra/d* (or *Notch1* and *Bcl11b*) in many instances it increased the association of the other genes with *Tcra/d* (**Figure 4B**). Moreover, in DP cells, *Tcra/d* was always closer to the other loci analyzed compared to *Igh*. It is interesting to note that even though *Bcl11b* is located adjacent to *Igh* on chromosome 12, the two loci associate with *Tcra/d* at different frequencies. In sum, these data indicate that RAG differentially influences the spatial relationship of RAG enriched loci at the DN and DP stages of development. The change in trend in the two cell types could be influenced by the enrichment of bound RAG, the level of transcription, and differences in the transcription factor binding profiles of the individual genes in these cells. What is clear though is that there is a shift in organization of these loci at the two different stage of development that in many cases is affected by the presence of RAG1.

### TRANSCRIPTION MEDIATED NUCLEOSOME RECONFIGURATION PROVIDES TRANSIENT RSS ACCESS TO RAG

Our data linking transcription with altered nuclear organization and targeted RAG cleavage of *Tcra* (11), go hand in hand with another recent paper that demonstrates the significance of transcription and altered chromatin organization in targeting of RAG-mediated breaks (24). Bevington and Boyes examined the requirements for activation of the *Igk* and *Igl* light chain loci by making use of interferon regulatory factor IRF4/IRF8 double deficient mice, which are blocked at the pro- to pre-B cell stage of development (25). One of the major defects of an absence of IRF4 is the loss of activation of enhancer elements on the *Igk* and *Igl* light chain loci, which are important for rearrangement (26). Restoration of light chain rearrangement and non-coding transcription can occur by enforced transgenic overexpression of IRF4 in *Irf4*<sup>-/-</sup>

*Irf8*<sup>-/-</sup> double mutant mice at the earlier pro-B cell stage of development. However, in contrast to wild-type pre-B cells (where the *Igk* locus is recombined before *Igl*), the *Igl* locus is rearranged in preference to *Igk* in the transgenic mice. This *in vivo* model enabled them to examine the mechanisms underlying RSS accessibility at these two loci. Surprisingly, they find that high levels of H3K4me3 at J $\kappa$  RSSs only leads to partial activation of recombination at the *Igk* locus indicating that enrichment of this mark is not tightly linked to recombination. Furthermore, they show that neither H3 and H4 histone acetylation is sufficient to increase RSS accessibility although enrichment of histone acetylation appears to be linked to early activation of the two loci. In addition they find no correlation between H3K36me3 enrichment and recombination at the *Igl* locus. Thus, none of the chromatin modifications that are associated with recombination are tightly linked to increased RSS accessibility. In contrast, increased accessibility, as measured by restriction enzyme digestion close to RSS sites in *Igl* versus *Igk*, correlates well with increased *Igl* recombination. Importantly, inhibition of transcription, through treatment with  $\alpha$ -amanitin, abrogates restriction enzyme accessibility.

To further explore the links between recombination, restriction enzyme digest efficiency and transcription Bevington and Boyes analyze changes in nucleosome composition as a potential mechanism for increasing RSS accessibility. Specifically they examine eviction of an H2A/H2B dimer from nucleosomes, which temporarily converts the latter to hexasomes during the passage of RNA polymerase (27–29). Nucleosome to hexasome conversion is known to reduce histone/DNA contacts and release 35–40 bp of DNA, which could permit RAG access to RSS sites. To test this prediction they analyzed uncoupled cleavage on *in vitro* substrates and show that RAG is more efficient in cleaving hexasomes than nucleosomes. Furthermore, they demonstrate that there is an inverse relationship between elevated levels of RNA polymerase II and the presence of H2A/H2B (see model in **Figure 5**). RSS accessibility is transient, which fits with the observed displacement time for H2A/H2B eviction of 6 min during passage of RNA polymerase II (27). In the context of recombination, transient RSS accessibility implies that only a limited number of RSSs will be available for recombination at any one time.

Bevington and Boyes speculate that the transient nature of RSS accessibility may be important to inhibit excess RAG cutting and genome instability. This fits well with our data showing that mono-allelic RAG cleavage occurs on the 3' end of looped out *Tcra* alleles, and that looping out occurs in only a subset of cells and thus it is likely to be a transient event (see model in **Figure 5**). Furthermore, we demonstrate that looping out is linked to high levels of transcription at the 3' end of *Tcra* and is dependent on the presence of RAG1. Feedback regulation of RAG cleavage by ATM and the C-terminus of RAG2 involve changes in nuclear accessibility that include inhibition of higher-order loop formation and repositioning of the uncleaved antigen-receptor allele to repressive PCH. In the absence of these changes we find an increase in bi-allelic and bi-locus cleavage and genome instability.

## CONCLUDING REMARKS

With the current systems that are available it is difficult to study the cascade of interdependent events that occur during V(D)J

recombination, particularly since many of these may be very transient. The challenge now is to develop live cell systems in which cleavage and repair can be analyzed over a period of time in single cells. This is no trivial matter when it comes to the antigen-receptor loci as these are spread over megabases of DNA. In addition to labeling of different regions along the loci, the RAG1/2 recombinase and components of the DNA damage response and repair machineries need to be tagged for visualization. Until such systems are in place the next best approach is to focus on single cell analyses.

## ACKNOWLEDGMENTS

This work is supported by the following grants: NIH R01GM086852 (Jane A. Skok and Julie Chaumeil), NIH R56NIAIDAI099111 (Jane A. Skok). Jane A. Skok is an LLS scholar. Julie Chaumeil is an Irvington Institute Fellow of the Cancer Research Institute.

## REFERENCES

- Johnson K, Chaumeil J, Skok JA. Epigenetic regulation of V(D)J recombination. *Essays Biochem* (2010) **48**:221–43. doi:10.1042/bse0480221
- Hewitt SL, Chaumeil J, Skok JA. Chromosome dynamics and the regulation of V(D)J recombination. *Immunol Rev* (2010) **237**:43–54. doi:10.1111/j.1600-065X.2010.00931.x
- Matthews AG, Kuo AJ, Ramon-Maiques S, Han S, Champagne KS, Ivanov D, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* (2007) **450**:1106–10. doi:10.1038/nature06431
- Liu Y, Subrahmanyam R, Chakraborty T, Sen R, Desiderio S. A plant homeodomain in RAG-2 that binds hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity* (2007) **27**:561–71. doi:10.1016/j.immuni.2007.09.005
- Ji Y, Resch W, Corbett E, Yamane A, Casellas R, Schatz DG. The *in vivo* pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell* (2010) **141**:419–31. doi:10.1016/j.cell.2010.03.010
- Lin YC, Benner C, Mansson R, Heinz S, Miyazaki K, Miyazaki M, et al. Global changes in the nuclear positioning of genes and intra- and interdomain genomic interactions that orchestrate B cell fate. *Nat Immunol* (2012) **13**:1196–204. doi:10.1038/ni.2432
- Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, Andrews S, et al. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet* (2010) **42**:53–61. doi:10.1038/ng.496
- Cheutin T, Cavalli G. Progressive polycomb assembly on H3K27me3 compartments generates polycomb bodies with developmentally regulated motion. *PLoS Genet* (2012) **8**:e1002465. doi:10.1371/journal.pgen.1002465
- Brandt VL, Hewitt SL, Skok JA. It takes two: communication between homologous alleles preserves genomic stability during V(D)J recombination. *Nucleus* (2010) **1**:23–9. doi:10.4161/nucl.1.1.10595
- Hewitt SL, Yin B, Ji Y, Chaumeil J, Marszalek K, Tenthorey J, et al. RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. *Nat Immunol* (2009) **10**:655–64. doi:10.1038/ni.1735
- Chaumeil J, Micsinai M, Ntziachristos P, Deriano L, Wang JM, Ji Y, et al. Higher-order looping and nuclear organization of *Tcra* facilitate targeted RAG cleavage and regulated rearrangement in recombination centers. *Cell Rep* (2013) **3**(2):359–70. doi:10.1016/j.celrep.2013.01.024
- Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* (2001) **2**:292–301. doi:10.1038/35066075
- Muller I, Boyle S, Singer RH, Bickmore WA, Chubb JR. Stable morphology, but dynamic internal reorganisation, of interphase human chromosomes in living cells. *PLoS One* (2010) **5**:e11560. doi:10.1371/journal.pone.0011560
- Chambeyron S, Bickmore WA. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* (2004) **18**:1119–30. doi:10.1101/gad.292104
- Kalhor R, Tjong H, Jayatilaka N, Alber F, Chen L. Genome architectures revealed by tethered chromosome conformation capture and population-based modeling. *Nat Biotechnol* (2012) **30**:90–8. doi:10.1038/nbt.2057

16. Boyle S, Rodesch MJ, Halvensleben HA, Jeddeloh JA, Bickmore WA. Fluorescence in situ hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis. *Chromosome Res* (2011) **19**:901–9. doi:10.1007/s10577-011-9245-0
17. Volpi EV, Chevret E, Jones T, Vatcheva R, Williamson J, Beck S, et al. Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *J Cell Sci* (2000) **113**(Pt 9):1565–76.
18. Deriano L, Chaumeil J, Coussens M, Multani A, Chou Y, Alekseyenko AV, et al. The RAG2 C terminus suppresses genomic instability and lymphomagenesis. *Nature* (2011) **471**:119–23. doi:10.1038/nature09755
19. Zha S, Bassing CH, Sanda T, Brush JW, Patel H, Goff PH, et al. ATM-deficient thymic lymphoma is associated with aberrant Tcrd rearrangement and gene amplification. *J Exp Med* (2010) **207**:1369–80. doi:10.1084/jem.20100285
20. Callen E, Bunting S, Huang CY, Difilippantonio MJ, Wong N, Khor B, et al. Chimeric IgH-TCRalpha/delta translocations in T lymphocytes mediated by RAG. *Cell Cycle* (2009) **8**:2408–12. doi:10.4161/cc.8.15.9085
21. Liyanage M, Weaver Z, Barlow C, Coleman A, Pankratz DG, Anderson S, et al. Abnormal rearrangement within the alpha/delta T-cell receptor locus in lymphomas from Atm-deficient mice. *Blood* (2000) **96**:1940–6.
22. Chaumeil J, Micsinai M, Ntziachristos P, Roth DB, Aifantis I, Kluger Y, et al. The RAG2 C-terminus and ATM protect genome integrity by controlling antigen receptor gene cleavage. *Nat Commun* (2013) **4**:2231. doi:10.1038/ncomms3231
23. Onozawa M, Aplan PD. Illegitimate V(D)J recombination involving nonantigen receptor loci in lymphoid malignancy. *Genes Chromosomes Cancer* (2012) **51**:525–35. doi:10.1002/gcc.21942
24. Bevington S, Boyes J. Transcription-coupled eviction of histones H2A/H2B governs V(D)J recombination. *EMBO J* (2013) **32**:1381–92. doi:10.1038/emboj.2013.42
25. Lu R, Medina KL, Lancki DW, Singh H. IRF-4,8 orchestrate the pre-B-to-B transition in lymphocyte development. *Genes Dev* (2003) **17**:1703–8. doi:10.1101/gad.1104803
26. Johnson K, Hashimshony T, Sawai CM, Pongubala JM, Skok JA, Aifantis I, et al. Regulation of immunoglobulin light-chain recombination by the transcription factor IRF-4 and the attenuation of interleukin-7 signaling. *Immunity* (2008) **28**(3):335–45. doi:10.1016/j.immuni.2007.12.019
27. Kimura H, Cook PR. Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. *J Cell Biol* (2001) **153**:1341–53. doi:10.1083/jcb.153.7.1341
28. Kireeva ML, Walter W, Tchernajenko V, Bondarenko V, Kashlev M, Studitsky VM. Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. *Mol Cell* (2002) **9**:541–52. doi:10.1016/S1097-2765(02)00472-0
29. Thiriet C, Hayes JJ. Histone dynamics during transcription: exchange of H2A/H2B dimers and H3/H4 tetramers during pol II elongation. *Results Probl Cell Differ* (2006) **41**:77–90. doi:10.1007/400\_009

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 October 2013; accepted: 20 November 2013; published online: 06 December 2013.

Citation: Chaumeil J and Skok JA (2013) A new take on V(D)J recombination: transcription driven nuclear and chromatin reorganization in RAG-mediated cleavage. *Front. Immunol.* **4**:423. doi: 10.3389/fimmu.2013.00423

This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2013 Chaumeil and Skok. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# CTCF and ncRNA regulate the three-dimensional structure of antigen receptor loci to facilitate V(D)J recombination

Nancy M. Choi and Ann J. Feeney\*

Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, CA, USA

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

Amy L. Kenter, University of Illinois College of Medicine, USA

Kay L. Medina, Mayo Clinic, USA

## \*Correspondence:

Ann J. Feeney, Department of Immunology and Microbial Science, The Scripps Research Institute, 10550 North Torrey Pines Road, IMM-22, La Jolla, CA 92037, USA  
e-mail: feeney@scripps.edu

At both the immunoglobulin heavy and kappa light chain loci, there are >100 functional variable (V) genes spread over >2 Mb that must move into close proximity in 3D space to the (D)J genes to create a diverse repertoire of antibodies. Similar events take place at the T cell receptor (TCR) loci to create a wide repertoire of TCRs. In this review, we will discuss the role of CTCF in forming rosette-like structures at the antigen receptor (AgR) loci, and the varied roles it plays in alternately facilitating and repressing V(D)J rearrangements. In addition, non-coding RNAs, also known as germline transcription, can shape the 3D configuration of the *Igh* locus, and presumably that of the other AgR loci. At the *Igh* locus, this could occur by gathering the regions being transcribed in the  $V_H$  locus into the same transcription factory where  $I\mu$  is being transcribed. Since the  $I\mu$  promoter,  $E\mu$ , is adjacent to the  $DJ_H$  rearrangement to which one V gene will ultimately rearrange, the process of germline transcription itself, prominent in the distal half of the  $V_H$  locus, may play an important and direct role in locus compaction. Finally, we will discuss the impact of the transcriptional and epigenetic landscape of the *Igh* locus on  $V_H$  gene rearrangement frequencies.

**Keywords: V(D)J recombination, antigen receptor, chromatin, non-coding RNA, CTCF, histone modification, chromatin loop**

## INTRODUCTION

Antigen receptor (AgR) loci are facing a uniquely difficult task to produce a great diversity of receptors in order to recognize the limitless possibility of antigens present in the environment of an organism. With the advent of next generation sequencing, we can now determine the actual diversity of AgRs by sequencing all of the rearrangements from developing B and T cells. This diversity is created through the combinatorial recombination of multiple variable (V), diversity (D), and joining (J) gene segments at AgR loci by the RAG1/2 recombinase complex, along with the extensive junctional diversity at the V–D, D–J, and V–J junctions.

One of the most extensively studied AgR loci is the mouse *Igh* locus where the  $V_H$ ,  $D_H$ , and  $J_H$  gene segments span a region of ~2.8 Mb (Figure 1). The 8–13  $D_H$  genes, the four  $J_H$  genes, and all of the constant region genes and enhancers are located within a relatively small 300 kb region. In contrast, the 195  $V_H$  genes, of which ~100 were deemed to be functional, are spread out over ~2.5 Mb. To create the greatest combinatorial diversity, all V genes would have to be able to access the  $D_H$  and  $J_H$  genes relatively equally regardless of their genomic distance. The question is then, how is this equality achieved?

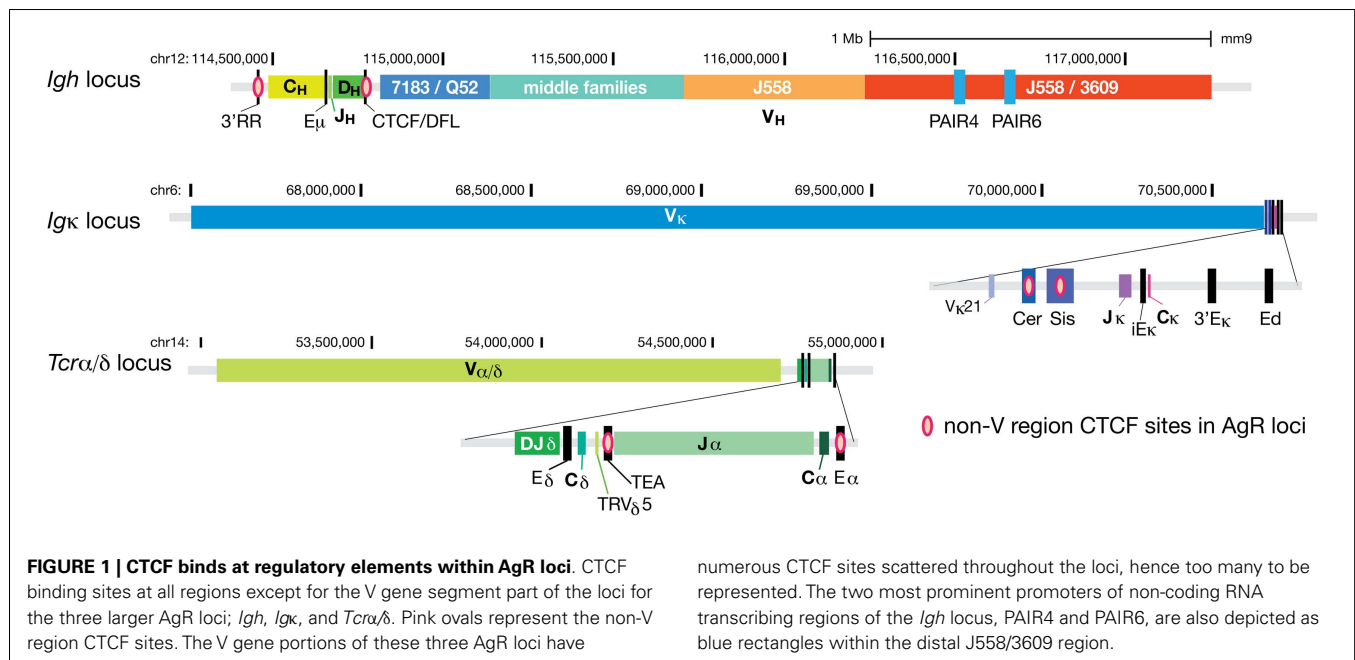
With growing appreciation for how three-dimensional structural changes at the locus may bring V genes into proximity of the (D)J rearrangement to which one V gene will ultimately rearrange, current studies are employing cutting edge technologies to further understand this process. Chromatin conformation capture (3C) and its more recent modifications, 4C, 5C, and Hi-C (1–3), have allowed the identification of long-range chromosomal interactions, which facilitate the rearrangement of distant V genes by

making critical connections between the V genes and enhancers downstream (4). Next generation sequencing technologies coupled with chromatin immunoprecipitation (ChIP) (ChIP-seq) have allowed us to determine the binding sites of transcription factors throughout the genome as well as the genome-wide epigenetic landscape. Deep sequencing of RNA reveals the entire transcriptional profile of cells for both coding and non-coding RNA (ncRNA). Together, these techniques supply us with a bounty of information regarding the transcriptional and epigenetic profile of AgR loci at varying stages of differentiation. In this review, we will summarize and discuss how these recent studies have advanced our understanding of how long-range chromatin interactions and epigenetic changes may regulate V(D)J recombination at mouse AgR loci.

## AgR LOCI UNDERGO LARGE SCALE THREE-DIMENSIONAL CONFORMATIONAL CHANGES DURING V(D)J REARRANGEMENT

All B cell and T cell receptor (BCR, TCR) subunits are formed through the process of V(D)J recombination. The BCR consists of two immunoglobulin heavy chains (*Igh*) and two identical light chains encoded by either the kappa (*Igk*) or lambda (*Igl*) loci. The TCR alpha ( $Tcr\alpha$ ) and beta ( $Tcr\beta$ ), or delta ( $Tcr\delta$ ) and gamma ( $Tcr\gamma$ ) chains constitute the TCR complex of the two major T cell subsets. The *Igh* and *Igk* are of similarly large sizes of approximately 2.8 and 3.2 Mb, while the  $Tcr\alpha/\delta$  and  $Tcr\beta$  loci are smaller at 1.7 and 0.66 Mb. In comparison, the *Igl* and *Tcr\gamma* loci are much smaller, each only being about 200 kb. The challenge, which is particularly great for the large receptor loci, is to give all V genes





a chance to undergo rearrangement in order to create a diverse repertoire. How an AgR locus brings the V genes into proximity to the (D)J genes to create this diversity is still an unanswered question.

The original observations that showed three-dimensional structural changes at the *Igh* locus, presumably facilitating the creation of a diverse AgR repertoire, came from fluorescent *in situ* hybridization (FISH) studies (5). It was found that the *Igh* and *Igk* loci were predominantly located at the periphery of the nucleus in non-recombining cell types, but were found in more centralized locations in B cells. The nuclear periphery is generally considered a transcriptionally silent environment and is associated with repressive chromatin modifications, whereas gene dense active regions of the genome are more centrally located (6). Using two colors of probes at proximal and distal ends of the *VH* locus, it was also shown for the first time that the *Igh* locus was in a more compacted conformation in recombining B cells. Subsequently, lineage- and developmental stage-specific locus contraction was observed for all of the large AgR loci: *Igk*, *Tcra/δ*, and *Tcrβ* (7–10). This process of locus contraction is reversible, as demonstrated by the extension of the *Igh* locus in pre-B cells, when *Igh* rearrangement is complete (7). Contraction and re-extension of the distal end of the *Tcra/δ* locus was also observed in double positive (DP) T cells (8). At this locus, contraction is necessary in double negative (DN) T cells for the accessibility of V genes used in TCRδ rearrangements, but in DP thymocytes, rearrangement of the more J-proximal *Vα* genes occurs before the rearrangement of distal *Vα* genes, so extension of the distal *Vα* genes would facilitate the ordered rearrangement of TCR *Vα* genes.

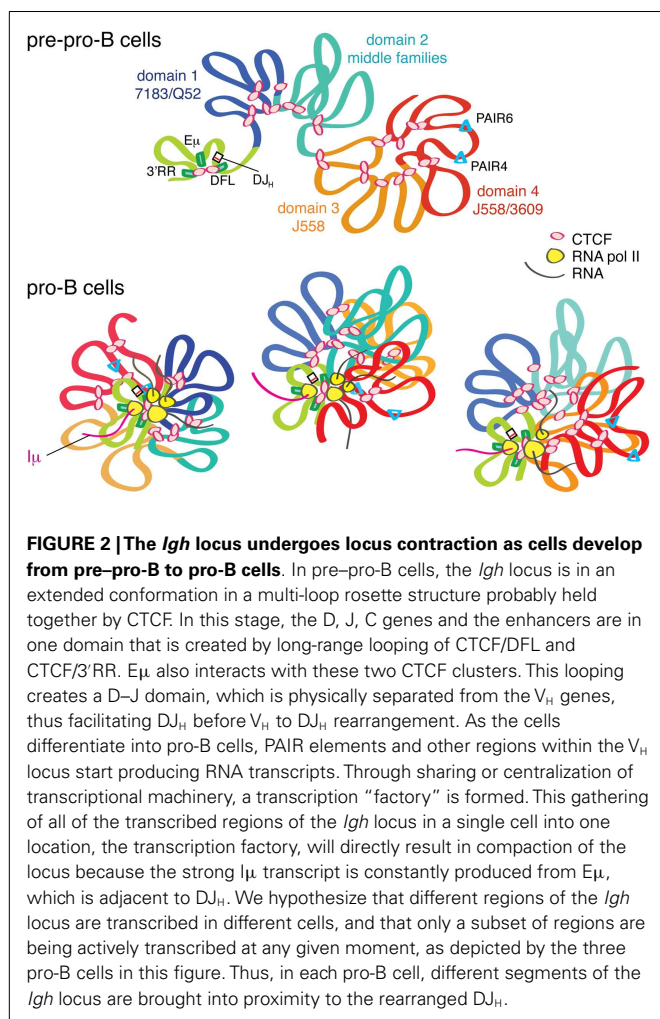
Greater insight to how such large-scale locus contraction may occur came from a 3D-FISH study by Jhunjhunwala et al. that used multiple 10 kb probes spanning the entire *Igh* locus followed by 3D computational reconstruction of the location of all the probe binding sites (11). The results showed that the locus could be divided

into three ~1 Mb compartments in pre-pro-B cells in which multiple chromatin loops formed rosette-like structures (Figure 2). These compartments then collapsed into a single globule as cells developed into pro-B cells. This brought the distal *VH* region into closer proximity within 3D space to the *DJH* genes and regulatory elements, and in fact the distal *VH* genes were found to be a similar distance away from the *DJH* region as the proximal *VH* genes (11).

It has been demonstrated that locus contraction of the *Igh* locus is regulated, directly or indirectly, by several key transcription factors. Mice deficient in YY1, Pax5, or the histone methyltransferase Ezh2 were impaired in locus contraction and in the rearrangement of distal *VH* genes (12–15). Ikaros has also been implicated in *Igh* locus contraction (16), but Rag1/2 is not required for this process (5). Together, these studies suggest that contraction is a pre-requisite state for efficient recombination of distal *VH* genes. Nonetheless, while AgR locus contraction is well established as a shared process among the large AgR loci that brings distal regions into closer 3D proximity to J genes prior to recombination, it has not been firmly determined what factors may be executing this task in the different lineages.

### CTCF AND COHESIN BIND EXTENSIVELY WITHIN AgR LOCI

CTCF is an 11 zinc-finger protein that is the only known insulator binding protein in vertebrates (17, 18). Insulators are genetic regions that prevent heterochromatin on one side of the insulator from spreading into the other side. They can also prevent against positional effect variegation, or varied expression of transgenes, depending upon the site of integration in relation to where the insulator is located. Some insulators also have enhancer-blocking activity, where an enhancer cannot activate a promoter when separated by an insulator. It is now known that insulators function through CTCF that creates long-range chromatin interactions by binding to other CTCF bound sites (19). In this way, a domain is created by these chromatin loops, and activity or inactivity of



the genes within the domain is insulated from the activity of neighboring domains. In fact, CTCF has been found to play a major role in the establishment of the higher order organization of chromosomes genome-wide, and it is found at the boundaries of topological domains in numerous Hi-C studies (20–22).

CTCF is aided in this domain-creating function by cohesin. Cohesin's only known function until a few years ago was to hold sister chromatids together during mitosis by forming a ring around the sister chromatids with its four protein subunits (23). Now it is well recognized that cohesin is bound to many active CTCF sites, and thought to reinforce the loops created by the long-range CTCF-CTCF binding (24–26).

Because of the capability of CTCF to form long-range loops, we hypothesized that if CTCF were present at many sites in the AgR loci, it may play a role in determining the 3D structure of the loci and could possibly even influence locus contraction. Thus, we performed ChIP-chip, and subsequently ChIP-seq, to demonstrate that indeed CTCF was bound at numerous sites within the Ig loci, and was therefore an excellent candidate for creating multiple long-range loops (27, 28). If CTCF also had an important role in locus contraction, then we would predict that it would only be bound to the *IgH* locus in pro-B cells, the stage at which the *IgH*

locus undergoes contraction. However, we found by ChIP/qPCR that CTCF had a similar pattern of binding in pre-B cells and even in thymocytes, showing that CTCF binding was not lineage- or stage-specific (28). However, widespread binding of CTCF within the *IgH* locus was not observed in fibroblasts, demonstrating that the binding was at least lymphoid-specific. We then analyzed the binding pattern of cohesin by performing a ChIP/qPCR for Rad21, one of the cohesin subunits. This revealed that the level of Rad21 binding was higher in pro-B cells than in pre-B cells or thymocytes for many sites, suggesting cohesin may have a greater role than CTCF in specifying the developmental stage in which *IgH* recombination occurs (28).

CTCF displayed more lineage- and developmental stage-specific binding at the *IgK* locus (28). Some sites were only bound in pre-B cells, while others showed lower levels of binding in pro-B cells or thymocytes. Rad21 binding also displayed similar lineage and stage-specificity at the *IgK* locus. Investigation of ChIP-seq of CTCF binding at the large TCR loci showed various extents of lineage- and stage-specificity. At all AgR loci, however, we observed that the binding of cohesin was highest in the appropriate lineage and developmental stage. From these observations, it can be seen that CTCF and Rad21 may have different degrees of function in regulating lineage and stage-specific 3D structures at each AgR locus.

### CTCF AND COHESIN INFLUENCE THE THREE-DIMENSIONAL STRUCTURE OF ANTIGEN RECEPTOR LOCI

To determine if CTCF made long-range loops that contributed to the compacted structure of the *IgH* locus in pro-B cells, we knocked down CTCF expression in RAG<sup>-/-</sup> pre-B cells that were cultured in IL7 for 4 days (27). 3D-FISH was performed 4 days after knock-down of CTCF, and the spatial distance between two probes at the far ends of the *IgH* locus did increase, although not to the extent observed in YY1-deficient pro-B cells. This could be due to the fact that while CTCF binding was significantly reduced it was not completely eliminated at the *IgH* locus in the knocked-down pro-B cells as determined by ChIP. However, it is likely that CTCF is only one of many factors that are involved in the compacted structure of the *IgH* locus.

Further insight into the contribution of CTCF to the 3D structure of the *IgH* locus came from the 4C studies of Guo et al. (4). They described two different kinds of loops that formed at the *IgH* locus: E $\mu$ -dependent and E $\mu$ -independent loops. Using a CTCF ChIP-loop assay, they showed that the proximal regions had several CTCF-dependent and E $\mu$ -independent interactions, spanning a region of ~140 kb, as well as interactions with CTCF/DFL. Using a probe in the distal J558 region in the CTCF ChIP-loop assay, they demonstrated four sites of interaction within a 500 kb region, about half of the number of sites seen in 4C with the same distal probe. Importantly, none of the distal CTCF-dependent loops interacted with any other part of the *IgH* locus, and similarly the loops in the proximal region only interacted locally. Jhunjhunwala et al. previously demonstrated that the *IgH* locus consisted of three distinct rosette-like multi-looped structures in pre-pro-B cells that compacted upon themselves during locus contraction (11). Thus, it may be that most of the CTCF-dependent loops described by Guo et al. are local interactions that form the basic rosette-like

loops within the *Igh* locus. In addition to CTCF-mediated loops, locus contraction results from further large-scale interactions of these rosettes that are dependent upon  $E\mu$ . It may be that the longer range interactions require other key transcription factors such as YY1 and Pax5. YY1 binds to  $E\mu$ , and Pax5 binds to PAIR elements, the sites of greatest antisense transcription (29, 30). Whether these are the regions of most importance for YY1 and Pax5 binding with regard to locus contraction, or whether their primary influence is indirect, is not known. Our previous results that showed an increase in spatial distance between the two ends of the *Igh* locus after CTCF knockdown may reflect a loosening of the individual rosette structures while still being held together by other locus contraction regulating factors.

### INSULATOR CTCF SITES BETWEEN THE V REGIONS AND D/J GENES AT AgR LOCI REGULATE REPERTOIRE DIVERSITY

The *Igh* locus has a pair of CTCF sites 3–5 kb upstream of the last functional  $D_H$  gene, DFL16.1 (28) (Figure 1). We and others have shown that this pair of CTCF sites (CTCF/DFL) has enhancer-blocking insulator activity in a traditional *in vitro* insulator assay (28, 31). By 3C, we have shown that CTCF/DFL loops to the cluster of nine CTCF sites downstream of the 3' regulatory region (3'RR) and to  $E\mu$  (27), and this was subsequently confirmed by two other groups (4, 32). Coincidentally, Jhunjhunwala et al. utilized a probe near CTCF/DFL in their trilateration study (11), so we know that this  $D_H$  and  $J_H$  gene containing-loop is located far from the  $V_H$  genes in pre-pro-B cells, but it moves in close proximity to  $V_H$  genes in pro-B cells (Figure 2). We hypothesized that this loop creates a domain that contains all the  $D_H$ ,  $J_H$  and constant region genes as well as the  $E\mu$  enhancer, but excludes  $V_H$  genes (27). This would provide a physical environment in which  $D_H$  to  $J_H$  rearrangement could occur without any  $V_H$  genes in the vicinity.

Since the  $D_H$  genes have much antisense transcription, it was hypothesized that perhaps the function of CTCF/DFL was to stop antisense transcription from extending into the proximal  $V_H$  genes, preventing accessibility of those  $V_H$  genes (31). Indeed, deletion of the entire 96 kb intervening region between DFL16.1 and 7183.2.3 resulted in increased levels of  $D_H$  antisense transcription and extension of this transcription into the proximal  $V_H$  locus (33). However, knockdown of CTCF in pro-B cells with an intact *Igh* locus only resulted in extension of the antisense transcription for ~4 kb, and the antisense transcription dropped precipitously at the 3' Adam6 gene (27). Thus, preventing  $D_H$  region antisense transcription from extending into the  $V_H$  region does not seem to be the function of CTCF/DFL.

Importantly, Guo et al. deleted or mutated the CTCF/DFL sites, and the consequences were profound (32). Ordered rearrangement was perturbed, such that  $V_H$  to  $D_H$  rearrangement occurred as well as  $D_H$  to  $J_H$  rearrangement. More strikingly, rearrangements were confined to the two most proximal  $V_H$  genes. This shows that one critical function of these CTCF/DFL sites is to allow the creation of a diverse repertoire of *Igh* rearrangement, fully utilizing all of the  $V_H$  genes, although the mechanism by which this is achieved is not clear (34). In addition to these striking changes, deletion of CTCF/DFL resulted in a lack of lineage restriction, with  $V_H$  rearrangement being observed in thymocytes. Thus, two of the basic tenets of the accessibility hypothesis, ordered rearrangements

and lineage- and stage-specific restriction of V(D)J rearrangement, are regulated by this pair of CTCF binding sites at CTCF/DFL.

The *Igk* locus has two pairs of CTCF sites between the  $V_k$  and  $J_k$  genes (28) (Figure 1). One pair is within a region called "Sis" (Silencer in the Intervening Sequence), which also contains several Ikaros binding sites (35). When Garrard and colleagues deleted the 650 bp Sis element in the germline (36), these mice showed a modest preference for rearranging proximal  $V_k$  over distal  $V_k$  genes, and sense non-coding transcription over  $V_k$  genes was also slightly increased. Much more striking was the germline deletion of the strong CTCF sites upstream of Sis in the region called "Cer" (Contracting Element for Recombination) (37). In the *Cer*<sup>-/-</sup> mice, sense transcription over a few proximal  $V_k$  genes was increased modestly, but there was a very strong bias toward rearrangement of the most proximal  $V_k$  genes and a great reduction of rearrangement of the remainder of genes. This effect was reminiscent of the strong over utilization of the most proximal  $V_H$  genes in the CTCF/DFL deletion mice (32). Significantly, some *Igk* rearrangement was observed in thymocytes in *Cer*<sup>-/-</sup> mice (although mainly limited to  $J_k1$ ), suggesting that the insulator sequences downstream of the V genes in both *Igh* and *Igk* loci are major contributors to the lineage restriction of Ig rearrangement. It should be mentioned that the *Igk* locus contraction was also reduced in *Cer*<sup>-/-</sup> mice, meaning extension of the locus could be a reason for the strong bias toward the most proximal V genes. Nonetheless, CTCF/DFL knockout mice did not display any change in *Igh* locus compaction (32), suggesting different modes of repertoire restriction at the two AgR loci.

In addition to the above studies in which the CTCF sites downstream of the V loci have been deleted, CTCF-deficient mice have been studied for their effects on repertoire formation. Hendriks and colleagues examined the *Igk* locus in mice carrying a B lineage-specific deletion of CTCF (38). By expressing a rearranged *Igh* gene they partially rescued development into pre-B cells. Absence of CTCF in pre-B cells resulted in a strong shift of usage to the most proximal  $V_k$  genes, where most rearrangements occurred at the 10 most proximal genes within the first ~200 kb in the knockout mice.  $V_k$  ncRNA were increased in this region, while the remainder of  $V_k$  ncRNA remained the same. Using Sis as an anchor/viewpoint for 4C-seq, it was demonstrated that the interactions of Sis with the 300 kb proximal region increased significantly. In contrast, iEk and 3'Ek viewpoints demonstrated that the enhancer interactions increased with sites up to 1 Mb into the  $V_k$  locus. However, other than a minor decrease of interaction of 3'Ek with the end of the  $V_k$  locus, the interactions of these three regulatory regions with the distal half of the  $V_k$  locus was unchanged. From these results, it seems that the majority of these long-range interactions between the enhancers or Sis with the distal 2/3 of the  $V_k$  locus are CTCF-independent interactions. Considering that the complete absence of CTCF in the cells gave a similar phenotype as the *Cer*<sup>-/-</sup> mice, the predominant effect of CTCF depletion throughout the *Igk* locus may be primarily due to the absence of CTCF binding at Cer.

As mentioned above, Rad21 (a subunit of cohesin) binds to CTCF sites in the AgR loci when rearrangement occurs (28, 39, 40). Seitan et al. analyzed the role of cohesin in V(D)J rearrangement at the *Tcr $\alpha$ / $\delta$*  locus (Figure 1) through the use of Rad21-deficient DP



thymocytes (39). Because cells cannot progress through cell division in the absence of cohesin, its role can only be ascertained in cells that do not divide, making DP thymocytes an appropriate cell type to study. They demonstrated that Rad21-deficiency resulted in reduced long-range looping between the CTCF/cohesin sites at TEA, the promoter of the germline transcripts of the 10 most 5' J $\alpha$  genes, and E $\alpha$  that also contains a CTCF/cohesin binding site. They also found an altered pattern of germline transcription in the J $\alpha$  region and reduced rearrangement to all but the most 5' J $\alpha$  genes in these Rad21-deficient mice.

A more detailed analysis of the role of CTCF/cohesin in TCR $\alpha$  rearrangement was performed using CTCF-deficient thymocytes (40). Shih et al. demonstrated by 3C that TEA and E $\alpha$  strongly interacted in wild type DP thymocytes, weakly in DN thymocytes, and not at all in B cells. TEA and E $\alpha$  also interacted with several proximal V $\alpha$  genes and with some J $\alpha$  genes, predominantly at the 5' portion of the J $\alpha$  region. In the Tcr $\alpha$ / $\delta$  locus, most functional V $\alpha$  genes have CTCF sites bound adjacent to the promoters, and thus it appears that normally CTCF nucleates a hub of proximal V $\alpha$  genes, a subset of J $\alpha$  genes, and the enhancer to create a functional recombination center. This entire hub of interactions was greatly reduced in E $\alpha$ -deficient DP thymocytes, and thus dependent upon E $\alpha$ . Deletion of TEA resulted in a shift of the peak of interaction of E $\alpha$  to the middle J $\alpha$  genes, likely explaining the previous observations that TEA deletion shifted the predominant rearrangements and germline transcription to the middle J $\alpha$  genes (41). In contrast to these results in wild type mice, 3C analysis of CTCF-deficient DP thymocytes revealed a reduction in the E $\alpha$  interactions with TEA, 5' J $\alpha$ , and certain 3' V $\alpha$  genes, and the level of rearrangement at the Tcr $\alpha$  locus was greatly reduced. Strikingly, the CTCF-deficient DP thymocytes showed increased E $\alpha$  contacts with the Tcr $\delta$  gene segments that are just upstream of TEA. Thus, it appears that the role of CTCF is to promote E $\alpha$  interactions with the 3' V $\alpha$  and 5' J $\alpha$  genes, while discouraging interactions with the intervening Tcr $\delta$  genes. 3D-FISH experiments demonstrated that the 3' end of the locus was still contracted in CTCF-deleted DP thymocytes, but 3C results showed that the long-range interactions were reduced for some 3' V $\alpha$  genes in DP thymocytes in the absence of CTCF. The level of transcription paralleled the new contacts as TEA-dependent transcription was decreased and transcription of Tcr $\delta$  genes was increased. Notably, this pattern of altered transcription and 3C contacts paralleled that seen in TEA<sup>-/-</sup> mice, suggesting that it is the CTCF binding to TEA in WT DP thymocytes that directs E $\alpha$  to interact with 5' V $\alpha$  and 3' J $\alpha$  and promotes their transcription and subsequent rearrangement. CTCF binding to TEA also presumably directs E $\alpha$  to skip over the more proximal Tcr $\delta$  genes and instead interact with the 5' V $\alpha$  genes further away in the locus. In this way, the function of the CTCF-binding region at TEA resembles that of CTCF/DFL and Cer/Sis in that it prevents interactions with the immediately proximal genes, and instead directs interactions to V genes that are further away, allowing the creation of a diverse repertoire of AgR.

### 3D CHANGES CAUSED BY NON-CODING RNA

For many years we have known that the J and C genes of each AgR locus undergo high levels of non-coding transcription when the

locus is undergoing rearrangement (42, 43). In addition, V genes can produce low levels of sense ncRNA (or “germline transcription”) when they are accessible for rearrangement (44). In a few cases it has been demonstrated that these sense ncRNAs begin at the V gene's promoter and stop shortly after the RSS and presumably this is the extent of most sense ncRNA. More recently, ncRNA in the antisense direction was described, and these ncRNAs are largely intergenic and longer (45). We performed directional RNA-seq of the *Igh* locus, thus defining all of the sense and antisense ncRNA within the locus in pro-B cells (29). Strikingly, there were three major regions of antisense ncRNA, and two minor antisense regions. The three major transcripts began at three of the PAIR elements, PAIR 4, 6, and 11. The 14 PAIR elements, or Pax5 Intergenic Repeat elements, consist of binding sites for Pax5, E2A, and CTCF. These regions have high levels of H3K4me3 and H3ac, as would be expected since they are so highly transcribed (29). The two minor regions of antisense ncRNA were in the proximal J558 region, the site of the originally described antisense RNA (45), and near the J606 genes.

It is now widely accepted that transcription takes place in sub-nuclear compartments called transcription factories, which are clusters of RNA polymerases (46, 47). Many genes are transcribed within each transcription factory, and often co-regulated genes occupy one together regardless of their genomic distance, and even genes on separate chromosomes may co-localize to the same factory (47, 48). It can be hypothesized that if all *Igh* ncRNA were to be transcribed from the same transcription factory, any regions within the V<sub>H</sub> part of the *Igh* locus that are being transcribed will of necessity be brought into juxtaposition with E $\mu$ , which contains the promoter of the predominant I $\mu$  germline transcript (29, 49). I $\mu$  is constantly transcribed and located 1–2.2 kb downstream of the J<sub>H</sub> genes (50). This would mean that any V<sub>H</sub> genes being transcribed would be close to the DJ<sub>H</sub> region to which one of the V<sub>H</sub> genes would ultimately rearrange in each pro-B cell (Figure 2). In support of this hypothesis, we demonstrated by 3C that PAIR4 and PAIR6, the regions of highest antisense transcription within the V<sub>H</sub> region, directly interacted with E $\mu$  (29). We knew that YY1<sup>-/-</sup> pro-B cells do not undergo locus contraction or rearrange distal V<sub>H</sub> genes. With this in mind, we also showed that YY1<sup>-/-</sup> pro-B cells did not undergo antisense transcription at PAIR elements, and their PAIR elements did not interact with E $\mu$  (29). Thus, it is possible that the lack of antisense ncRNA in the distal V<sub>H</sub> region of YY1<sup>-/-</sup> pro-B cells contributes to their lack of both locus contraction and rearrangement of distal J558 genes. We also saw a modest increase in antisense transcription at PAIR elements in CTCF-knockdown in RAG<sup>-/-</sup> pro-B cells, and 3C analysis showed modestly increased interactions of PAIR and E $\mu$ . This is consistent with the idea that these interactions are taking place in a common transcription factory (27). By 3D-FISH, larger spatial distances between the proximal and distal ends of the *Igh* locus were seen in pro-B cells with CTCF knockdown, suggesting that CTCF is likely assisting in forming multiple loops within the *Igh* locus that “loosen” as its expression is reduced. However, the increase in PAIR–E $\mu$  interactions that we observed with loss of CTCF expression suggests that CTCF is not a major player in the pro-B specific locus contraction process.

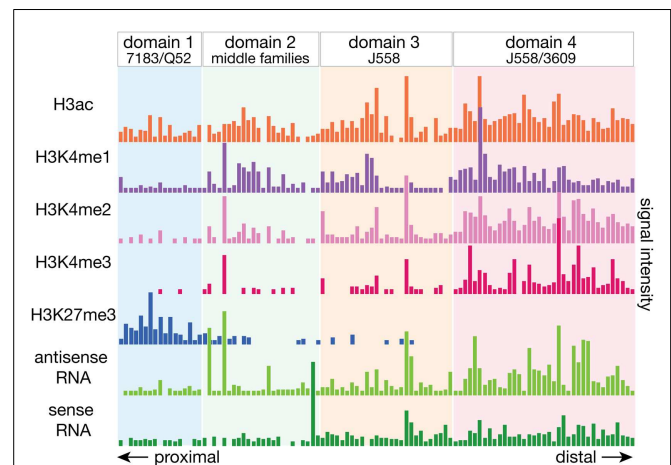
## DEEP SEQUENCING OF THE *Igh* REPERTOIRE IN PRO-B CELLS AND BIOINFORMATIC ANALYSES

While it is necessary to understand the effect of individual elements that regulate accessibility and chromatin structure at AgR loci, it is likely that many different factors are acting in concert for efficient production of a diverse repertoire. Recently, our lab and the Oltz lab adopted a bioinformatic approach with a goal to assign weight to the various factors that influence the frequency of rearrangement of individual V genes. To address this aim, we correlated the sequenced repertoires of mouse *Igh* and *Tcrβ* to ChIP-seq data for histone modifications and transcription factor binding and RNA-seq data for ncRNA transcripts (51, 52).

For the analysis of the mouse *Igh* repertoire in C57BL/6 mice, we sequenced 5'RACE-amplified cDNA from cell sorter purified pro-B cells to determine the pre-selection repertoire (51). Because this approach utilizes universal sequences to the 5' annealed adapter and Cμ on the expressed heavy chain transcript, it allows for an unbiased amplification of the expressed repertoire. In pro-B cells, as expected, the V<sub>H</sub> genes were recombined at widely different frequencies throughout the locus. We assessed the histone post-translational modifications and transcript levels over each actively recombined gene and observed a significant distinction between V<sub>H</sub> genes at the distal and proximal parts of the locus (Figure 3). Distal J558 family genes had greater enrichment for the active histone modifications (H3K4 methylation and H3 acetylation) as well as higher levels of both sense and antisense transcripts, than the proximal 7183 and Q52 families. This difference in epigenetic profiles suggests that these factors may be preferentially more influential at the distal half of the large *Igh* locus. We therefore divided the *Igh* locus into four domains based on V<sub>H</sub> gene family locations, and found that domain 1, consisting of the 7183 and Q52 families, had very low levels of H3K4 methylation and the lowest levels of ncRNA. Domain 4, the most distal, containing all of the 3609 family as well as half of the J558 genes, had the highest levels of all the active histone modifications as well as the highest levels of both sense and antisense ncRNA. Domain 3, containing the remainder of the J558 genes, also had active chromatin marks and higher levels of ncRNA than the proximal genes.

When the relation to CTCF and Rad 21 binding was examined, all but one actively utilized gene of the proximal 7183 and Q52 families in domain 1 had a CTCF binding site within 100 bp, and all but one inactive gene had a CTCF site at ~1–20 kb distance. While at a genomic scale, a distance of 100 bp vs. >1 kb may not be of great difference, it may be enough distinction to place an RSS in close enough vicinity to the recombination center at the J<sub>H</sub> region to provide a significant advantage to a V<sub>H</sub> gene. CTCF binding at the base of the loop at CTCF/DFL, which is proximal to the rearranged DJ<sub>H</sub>, and the base of the loop of functional V<sub>H</sub>-adjacent CTCF sites in domain 1 would bring these regions in close proximity. Genes in the middle and distal regions did not show this tendency, suggesting that having a close CTCF binding site is most important for the genes at the proximal end of the *Igh* locus.

We previously demonstrated that RSS quality could influence V<sub>H</sub> gene rearrangement frequency, and demonstrated that three different prototypic 7183 RSSs and a S107 RSS were more effective than a J558 RSS (53). All of the J558 RSSs are much further from the consensus RSS sequence than the 7183 RSSs. However, we also



**FIGURE 3 | The *Igh* locus can be divided into four domains by the epigenetic and transcriptional landscape.** The local epigenetic and transcriptional environment of each gene is plotted, with the numbers deriving from the total number of ChIP-seq or RNA-seq reads for the 2.5 kb region centered around each V<sub>H</sub> gene. Active histone modifications and ncRNA transcripts were enriched at V<sub>H</sub> genes at the distal end of the locus while proximal genes had very little of these features. Domains were divided by the boundary of V<sub>H</sub> gene families, and bioinformatic analyses of the various epigenetic elements suggest that genes in each domain may be regulated by different mechanisms.

showed that other parameters can override this effect, and that V genes with an identical RSS can rearrange at very different frequencies *in vivo* (53–55). Results from a computational model-building algorithm using our ChIP-seq, RNA-seq, and *Igh* repertoire deep sequencing data determined that having a functional RSS and an open chromatin environment as assessed by histone modifications were significant factors in predicting the activity of a V<sub>H</sub> gene (51). When just the actively rearranging functional V<sub>H</sub> genes were considered, the different domains of the V<sub>H</sub> locus had different factors that correlated with recombination frequency. Within the proximal domain 1, proximity to the DJ<sub>H</sub> genes was most significant, which is in agreement with the data we obtained a decade ago on another *Igh* haplotype, *Igh<sup>a</sup>*, in pro-B cells from μMT mice (53). In contrast, at the distal domains, higher levels of active histone modifications appeared to be most important. This greater enrichment for active histone modifications at the distal V<sub>H</sub> genes may reflect recruitment of these genes to the recombination center via transcription or some unknown factor that compensates for the disadvantages such as the distance from the DJ<sub>H</sub> genes and their poorer RSSs.

At the *Tcrβ* locus, Gopalakrishnan et al. took a different approach of assessing individual Vβ gene usage by using a Taqman assay to measure rearrangement of genomic DNA rather than the 5'RACE approach that we used for the *Igh* repertoire (52). This approach is feasible at the *Tcrβ* locus due to the much smaller number of V genes compared to the *Igh* locus. When recombination frequency was compared to 3C interaction data, there was no rearrangement advantage observed for Vβ genes that displayed higher levels of interaction with the Dβ1 gene, leading authors to conclude that once the contraction has occurred at the relatively

smaller *Tcrβ* locus, spatial access is not a determining factor for V $\beta$  gene usage. However, it should be noted that all but two of the V $\beta$  genes are present within 235 kb at this locus, whereas the *Igh* and *Igk* V genes are spread over >2.5 kb. Therefore, proximity of V genes to (D)J genes in 3D space is much more likely to contribute to V gene rearrangement frequency in the large *Igh* and *Igk* loci. The bioinformatic analysis of all of the chromatin modifications, transcriptional activity, and 3D proximity for the *Tcrβ* locus led to the conclusion that having a functional RSS, higher nucleosome depletion (FAIRE assay), and higher RNA pol II binding were good indicators for active vs. inert V $\beta$  genes. They also concluded, for the actively rearranging genes, higher levels of active histone modifications correlated with higher levels of recombination, similarly to our conclusions for the domain 3 and 4 V $H$  genes.

The results from the *Tcrβ* and *Igh* locus considered together suggest that while generally accessible chromatin conformation and functional RSS sequences are both important, the different AgR loci are not governed by the same rules. In the case of the *Igh* locus, even the proximal and distal ends of the locus may be regulated by different mechanisms, which is likely due to its great expansion over a large genomic area and hence a greater need for locus contraction to bring the distal and middle V $H$  genes closer.

### MODEL FOR THE ROLE OF CTCF AND ncRNA IN THE ESTABLISHMENT OF THE 3D STRUCTURE OF THE AgR LOCI

CTCF and its partner cohesin play important structural roles in creating large domains throughout the entire genome. Within AgR loci, there is a much higher density of CTCF/cohesin sites at rearranging loci than elsewhere in the genome. We hypothesize that the many CTCF/cohesin sites are necessary to create the multi-looped rosette-like structure that is the basic conformation of all AgR loci. This rosette structure makes it easier to compact various loci at the time of rearrangement. For some V genes, such as the V $H$  genes in domain 1 of the *Igh* locus, having a CTCF site near the RSS appears to be critical for a V $H$  gene to undergo rearrangement, but these V $H$  genes are rather poor in active histone marks and ncRNA. Thus, in lieu of these accessibility factors, being physically tethered to the recombination center, presumably by interactions with CTCF/DFL, is of great importance. In addition to the many CTCF sites throughout the large V gene portions of the AgR loci, CTCF/cohesin sites in between the V and J regions of the large AgR loci seem to be particularly important in regulating proper V gene rearrangements in a lineage- and developmental stage-specific manner (Figure 1). We also propose that ncRNA, or germline transcription, can directly facilitate *Igh* locus compaction if V $H$  genes or intergenic regions being transcribed are located in the same transcription factory as the  $I\mu$  ncRNA. Since the DJ $H$  rearrangement is directly adjacent to the highly transcribed  $I\mu$ , transcription will place the DJ $H$  rearrangement very close to the transcription factory. We hypothesize that the structure of the *Igh* locus is very dynamic in pro-B cells, with different subsets of V $H$  genes being transcribed in each pro-B cell (Figure 2, bottom). Thus, we suggest that the dynamic and stochastic nature of germline transcription will physically move different parts of the V $H$  gene locus into proximity to the DJ $H$  rearrangement in each pro-B cell, and this will provide equal opportunity for V $H$  genes throughout the locus to come into proximity to the DJ $H$  rearrangement. Presumably, this

same activity could take place at the other AgR loci. In this way, the production of diverse repertoires of antibodies and TCR is assured.

### ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health grants R01AI08218 and R21AI1007343.

### REFERENCES

- Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. *Science* (2002) **295**:1306–11. doi:10.1126/science.1067799
- Ohlsson R, Gondor A. The 4C technique: the ‘Rosetta stone’ for genome biology in 3D? *Curr Opin Cell Biol* (2007) **19**:321–5. doi:10.1016/j.ccb.2007.04.008
- Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* (2009) **326**:289–93. doi:10.1126/science.1181369
- Guo C, Gerasimova T, Hao H, Ivanova I, Chakraborty T, Selimyan R, et al. Two forms of loops generate the chromatin conformation of the immunoglobulin heavy-chain gene locus. *Cell* (2011) **147**:332–43. doi:10.1016/j.cell.2011.08.049
- Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* (2002) **296**:158–62. doi:10.1126/science.1068768
- Kind J, van Steensel B. Genome-nuclear lamina interactions and gene regulation. *Curr Opin Cell Biol* (2010) **22**:320–5. doi:10.1016/j.ccb.2010.04.002
- Roldan E, Fuxa M, Chong W, Martinez D, Novatchkova M, Busslinger M, et al. Locus ‘decontraction’ and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nat Immunol* (2005) **6**:31–41. doi:10.1038/ni1150
- Shih HY, Krangel MS. Distinct contracted conformations of the *Tcrα*/*Tcrδ* locus during *Tcrα* and *Tcrδ* recombination. *J Exp Med* (2010) **207**:1835–41. doi:10.1084/jem.20100772
- Sayegh C, Jhunjhunwala S, Riblet R, Murre C. Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. *Genes Dev* (2005) **19**:322–7. doi:10.1101/gad.1254305
- Skok JA, Gisler R, Novatchkova M, Farmer D, de Laat W, Busslinger M. Reversible contraction by looping of the *Tcrα* and *Tcrβ* loci in rearranging thymocytes. *Nat Immunol* (2007) **8**:378–87. doi:10.1038/ni1448
- Jhunjhunwala S, van Zelm MC, Peak MM, Cutchin S, Riblet R, van Dongen JJ, et al. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. *Cell* (2008) **133**:265–79. doi:10.1016/j.cell.2008.03.024
- Liu H, Schmidt-Suppran M, Shi Y, Hobeika E, Barteneva N, Jumaa H, et al. Yin Yang 1 is a critical regulator of B-cell development. *Genes Dev* (2007) **21**:1179–89. doi:10.1101/gad.1529307
- Hesslein DG, Pflugh DL, Chowdhury D, Bothwell AL, Sen R, Schatz DG. Pax5 is required for recombination of transcribed, acetylated, 5' IgH V gene segments. *Genes Dev* (2003) **17**:37–42. doi:10.1101/gad.1031403
- Fuxa M, Skok J, Souabni A, Salvagiotto G, Roldan E, Busslinger M. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev* (2004) **18**:411–22. doi:10.1101/gad.291504
- Su IH, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT, et al. Ezh2 controls B cell development through histone H3 methylation and IgH rearrangement. *Nat Immunol* (2003) **4**:124–31. doi:10.1038/ni876
- Reynaud D, Demarco IA, Reddy KL, Schjerve H, Bertolino E, Chen Z, et al. Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat Immunol* (2008) **9**:927–36. doi:10.1038/ni.1626
- Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell* (2009) **137**:1194–211. doi:10.1016/j.cell.2009.06.001
- Wallace JA, Felsenfeld G. We gather together: insulators and genome organization. *Curr Opin Genet Dev* (2007) **17**:400–7. doi:10.1016/j.gde.2007.08.005
- Phillips-Cremins JE, Corces VG. Chromatin insulators: linking genome organization to cellular function. *Mol Cell* (2013) **50**:461–74. doi:10.1016/j.molcel.2013.04.018
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* (2012) **485**:376–80. doi:10.1038/nature11082

21. Hou C, Li L, Qin ZS, Corces VG. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol Cell* (2012) **48**:471–84. doi:10.1016/j.molcel.2012.08.031
22. Botta M, Haider S, Leung IX, Lio P, Mozziconacci J. Intra- and inter-chromosomal interactions correlate with CTCF binding genome wide. *Mol Syst Biol* (2010) **6**:426. doi:10.1038/msb.2010.79
23. Nasmyth K, Haering CH. Cohesin: its roles and mechanisms. *Annu Rev Genet* (2009) **43**:525–58. doi:10.1146/annurev-genet-102108-134233
24. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* (2008) **132**:422–33. doi:10.1016/j.cell.2008.01.011
25. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, Schirghuber E, et al. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* (2008) **451**:796–801. doi:10.1038/nature06634
26. Rubio ED, Reiss DJ, Welsh PL, Distech CM, Filippova GN, Baliga NS, et al. CTCF physically links cohesin to chromatin. *Proc Natl Acad Sci U S A* (2008) **105**:8309–14. doi:10.1073/pnas.0801273105
27. Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, et al. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the *Igh* locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci U S A* (2011) **108**:9566–71. doi:10.1073/pnas.1019391108
28. Degner SC, Wong TP, Jankevicius G, Feeney AJ. Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J Immunol* (2009) **182**:44–8.
29. Verma-Gaur J, Torkamani A, Schaffer L, Head SR, Schork NJ, Feeney AJ. Non-coding transcription within the *Igh* distal  $V_H$  region at PAIR elements affects the 3D structure of the *Igh* locus in pro-B cells. *Proc Natl Acad Sci U S A* (2012) **109**:17004–9. doi:10.1073/pnas.1208398109
30. Ebert A, McManus S, Tagoh H, Medvedovic J, Salvagiotto G, Novatchkova M, et al. The distal  $V_H$  gene cluster of the *Igh* locus contains distinct regulatory elements with Pax5 transcription factor-dependent activity in pro-B cells. *Immunity* (2011) **34**:175–87. doi:10.1016/j.immuni.2011.02.005
31. Featherstone K, Wood AL, Bowen AJ, Corcoran AE. The mouse immunoglobulin heavy chain V-D intergenic sequence contains insulators that may regulate ordered V(D)J recombination. *J Biol Chem* (2010) **285**:9327–38. doi:10.1074/jbc.M109.098251
32. Guo C, Yoon HS, Franklin A, Jain S, Ebert A, Cheng HL, et al. CTCF-binding elements mediate control of V(D)J recombination. *Nature* (2011) **477**:424–30. doi:10.1038/nature10495
33. Giallourakis CC, Franklin A, Guo C, Cheng HL, Yoon HS, Gallagher M, et al. Elements between the *Igh* variable (V) and diversity (D) clusters influence antisense transcription and lineage-specific V(D)J recombination. *Proc Natl Acad Sci U S A* (2010) **107**:22207–12. doi:10.1073/pnas.1015954107
34. Feeney AJ, Verma-Gaur J. CTCF-cohesin complex: architect of chromatin structure regulates V(D)J rearrangement. *Cell Res* (2012) **22**:280–2. doi:10.1038/cr.2011.188
35. Liu Z, Widlak P, Zou Y, Xiao F, Oh M, Li S, et al. A recombination silencer that specifies heterochromatin positioning and Ikaros association in the immunoglobulin kappa locus. *Immunity* (2006) **24**:405–15. doi:10.1016/j.immuni.2006.02.001
36. Xiang Y, Zhou X, Hewitt SL, Skok JA, Garrard WT. A multifunctional element in the mouse *Igk* locus that specifies repertoire and *Ig* loci subnuclear location. *J Immunol* (2011) **186**:5356–66. doi:10.4049/jimmunol.1003794
37. Xiang Y, Park SK, Garrard WT.  $V_k$  gene repertoire and locus contraction are specified by critical DNase I hypersensitive sites within the  $V_k$ - $J_k$  intervening region. *J Immunol* (2013) **190**:1819–26. doi:10.4049/jimmunol.1203127
38. Ribeiro de Almeida C, Stadhouders R, de Bruijn MJ, Bergen IM, Thongjuea S, Lenhard B, et al. The DNA-binding protein CTCF limits proximal  $V_k$  recombination and restricts enhancer interactions to the immunoglobulin kappa light chain locus. *Immunity* (2011) **35**:501–13. doi:10.1016/j.immuni.2011.07.014
39. Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnoli T, Mira-Bontenbal H, Brown KE, et al. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* (2011) **476**:467–71. doi:10.1038/nature10312
40. Shih HY, Verma-Gaur J, Torkamani A, Feeney AJ, Galjart N, Krangel MS.  $Tcr\alpha$  gene recombination is supported by a  $Tcr\alpha$  enhancer- and CTCF-dependent chromatin hub. *Proc Natl Acad Sci U S A* (2012) **109**:E3493–502. doi:10.1073/pnas.1214131109
41. Villey I, Caillol D, Selz F, Ferrier P, de Villartay J-P. Defect in rearrangement of the most 5'  $Tcr\alpha$  following targeted deletion of T early a (TEA): implications for  $Tcr\alpha$  locus accessibility. *Immunity* (1996) **5**:331–42. doi:10.1016/S1074-7613(00)80259-9
42. Yancopoulos GD, Alt FW. Regulation of the assembly and expression of variable-region genes. *Ann Rev Immunol* (1986) **4**:339–68. doi:10.1146/annurev.iy.04.040186.002011
43. Lennon GG, Perry RP. C mu-containing transcripts initiate heterogeneously within the *IgH* enhancer region and contain a novel 5'-nontranslatable exon. *Nature* (1985) **318**:475–8. doi:10.1038/318475a0
44. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged  $V_H$  gene segments. *Cell* (1985) **40**:271–81. doi:10.1016/0092-8674(85)90141-2
45. Bolland DJ, Wood AL, Johnston CM, Bunting SF, Morgan G, Chakalova L, et al. Antisense intergenic transcription in V(D)J recombination. *Nat Immunol* (2004) **5**:630–7. doi:10.1038/ni1068
46. Mitchell JA, Fraser P. Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev* (2008) **22**:20–5. doi:10.1101/gad.454008
47. Cook PR. A model for all genomes: the role of transcription factories. *J Mol Biol* (2010) **395**:1–10. doi:10.1016/j.jmb.2009.10.031
48. Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, Andrews S, et al. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet* (2010) **42**:53–61. doi:10.1038/ng.496
49. Matheson LS, Corcoran AE. Local and global epigenetic regulation of V(D)J recombination. *Curr Top Microbiol Immunol* (2012) **356**:65–89. doi:10.1007/82\_2011\_137
50. Osborne CS, Chakalova L, Mitchell JA, Horton A, Wood AL, Bolland DJ, et al. Myc dynamically and preferentially relocates to a transcription factory occupied by *Igh*. *PLoS Biol* (2007) **5**:e192. doi:10.1371/journal.pbio.0050192
51. Choi NM, Loguercio S, Verma-Gaur J, Degner SC, Torkamani A, Su AI, et al. Deep sequencing of the murine *Igh* repertoire reveals complex regulation of nonrandom V gene rearrangement frequencies. *J Immunol* (2013) **191**:2393–402. doi:10.4049/jimmunol.1301279
52. Gopalakrishnan S, Majumder K, Predeus A, Huang Y, Koues OI, Verma-Gaur J, et al. Unifying model for molecular determinants of the preselection  $V\beta$  repertoire. *Proc Natl Acad Sci U S A* (2013) **110**:E3206–15. doi:10.1073/pnas.1304048110
53. Williams GS, Martinez A, Montalbano A, Tang A, Mauhar A, Ogwaro KM, et al. Unequal  $V_H$  gene rearrangement frequency within the large  $V_H$ 7183 gene family is not due to RSS variation, and mapping of the genes shows a bias of rearrangement based on chromosomal location. *J Immunol* (2001) **167**:257–63.
54. Feeney AJ. Genetic and epigenetic control of V gene rearrangement frequency. *Adv Exp Med Biol* (2009) **650**:73–81. doi:10.1007/978-1-4419-0296-2\_6
55. Feeney AJ, Atkinson MJ, Cowan MJ, Escuro G, Lugo G. A defective  $V_kA2$  allele in Navajos which may play a role in increased susceptibility to *Haemophilus influenzae* type b disease. *J Clin Invest* (1996) **97**:2277–82. doi:10.1172/JCI118669

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 20 December 2013; paper pending published: 07 January 2014; accepted: 28 January 2014; published online: 11 February 2014.

Citation: Choi NM and Feeney AJ (2014) CTCF and ncRNA regulate the three-dimensional structure of antigen receptor loci to facilitate V(D)J recombination. *Front. Immunol.* 5:49. doi: 10.3389/fimmu.2014.00049

This article was submitted to *B Cell Biology*, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Choi and Feeney. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Ubiquitination events that regulate recombination of immunoglobulin loci gene segments

Jaime Chao, Gerson Rothschild and Uttiya Basu\*

Department of Microbiology and Immunology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

Wenxia Song, University of Maryland, USA

Paolo Casali, University of Texas Health Science Center, USA

## \*Correspondence:

Uttiya Basu, Department of Microbiology and Immunology, College of Physicians and Surgeons, Columbia University, 701 West 168th Street, New York, NY 10032, USA  
e-mail: ub2121@columbia.edu

Programed DNA mutagenesis events in the immunoglobulin (Ig) loci of developing B cells utilize the common and conserved mechanism of protein ubiquitination for subsequent proteasomal degradation to generate the required antigen-receptor diversity. Recombinase proteins RAG1 and RAG2, necessary for V(D)J recombination, and activation-induced cytidine deaminase, an essential mutator protein for catalyzing class switch recombination and somatic hypermutation, are regulated by various ubiquitination events that affect protein stability and activity. Programed DNA breaks in the Ig loci can be identified by various components of DNA repair pathways, also regulated by protein ubiquitination. Errors in the ubiquitination pathways for any of the DNA double-strand break repair proteins can lead to inefficient recombination and repair events, resulting in a compromised adaptive immune system or development of cancer.

**Keywords:** RAG proteins, AID, V(D)J recombination, class switch recombination, somatic hypermutation, ubiquitination, DNA repair

## INTRODUCTION

B cells are developmentally programed to undergo DNA double-strand breaks (DSBs) in the immunoglobulin (Ig) locus as they generate the antibody diversity required for adaptive immunity. Immature B cells reside in the bone marrow and undergo V(D)J recombination, using the DNA endonuclease activity of RAG1 and RAG2, to rearrange multiple gene segments and select a V(D)J exon in the immunoglobulin heavy chain (IgH) and a VJ exon in the immunoglobulin light chain (IgL) loci (1–3). Following V(D)J rearrangement, IgM<sup>+</sup> B cells traverse to secondary organs (e.g., Peyer's patches or the spleen) to undergo two additional DNA alteration events, namely class switch recombination (CSR) and somatic hypermutation (SHM), using the activity of the DNA cytidine deaminase activation-induced cytidine deaminase (AID) (4). CSR is a chromosomal rearrangement–deletion event requiring single-strand breaks in close proximity to each other on both DNA strands for selection of the particular antibody isotype the B cell will produce (5). SHM is the incorporation of point mutations in the recombined V(D)J exon of the IgH and IgL loci to increase the affinity of the expressed antibody for its cognate antigen (6). In this review, we discuss how DNA mutagenesis during V(D)J recombination, CSR, and SHM is regulated by various protein ubiquitination events.

Ubiquitination is a post-translation modification utilized as a regulatory mechanism by cells. The process involves the sequential actions of an ubiquitin-activating enzyme (E1), an ubiquitin-conjugating enzyme (E2), and an ubiquitin ligase (E3) to covalently attach ubiquitin to a lysine residue of the target protein. Deemed initially to be merely a system to mark proteins for degradation by the 26S proteasome, the diversity of ubiquitination and downstream events has recently been increasingly appreciated, reinforcing the possibility of an “ubiquitin code” as another layer in the multifaceted landscape of molecular regulation [reviewed in Ref. (7)].

## RAG PROTEINS AND V(D)J RECOMBINATION

V(D)J recombination occurs in the IgH and IgL (comprised of Ig $\kappa$  and Ig $\lambda$ ) chains of B cells. At the IgH locus, V(D)J recombination first connects a diversity (D) to a joining (J) segment to form a coding and a signal joint, followed by a second recombination event to bring together a variable (V) segment to the preformed DJ segment. IgL loci, on the other hand, do not contain D segments and are therefore subject to only one recombination event to form a VJ coding segment. For T cells,  $\beta$  and  $\delta$  TCR loci parallel the IgH locus, first joining DJ segments before recombining to a V segment;  $\alpha$  and  $\gamma$  TCR are analogous to IgL loci (1–3).

Recombination activating genes 1 and 2, encoding proteins RAG1 and RAG2 respectively, are necessary and sufficient for the breaks and rearrangements during V(D)J recombination. RAG1 and RAG2, collectively referred to as RAG in this review, are lymphoid-specific proteins that cleave and join DNA segments during V(D)J recombination [reviewed in Ref. (8)]. The RAG proteins specifically recognize recombination signal sequences (RSSs) that flank each V, D, and J segment. RSSs are composed of conserved heptamer and nonamer sequences separated by a non-conserved gap of either 12 or 23 base pairs, named 12RSS and 23RSS, respectively. Upon recognition and binding of a 12RSS or 23RSS, the RAG proteins form a complex that then captures the alternate 12RSS/RAG or 23RSS/RAG complex for synapsis to form a paired complex [reviewed in Ref. (9)]. Synapsis occurs exclusively between different RSS for efficient recombination, known as the “12–23 rule.” After the paired complex is formed, the RAG proteins nick and cleave 5' of the RSS to produce a hairpin-closed coding end and a blunt signal end. Subsequent processing by non-homologous end joining (NHEJ) factors results in the final coding and signal joints, completing the recombination of a D to a J segment, a V to a DJ, or a V to a J segment. Signal joints containing the RSSs are excised, leaving coding joints as the operative DNA in the cell. How RAG1 and RAG2 interact with each other and

in what exact sequence remains to be determined. However, the inherent ability of RAG to cause DNA breaks and recombination is of concern, particularly if lesions occur outside of warranted loci, namely IgH, IgL, or the TCR, engendering the possibility of translocations with oncogenes and/or transformation of the cell to a malignant state. Therefore, effective regulation of RAG1 and RAG2 is crucial both for efficient V(D)J recombination to generate the diversity in the adaptive immune system and for avoidance of genomic instability.

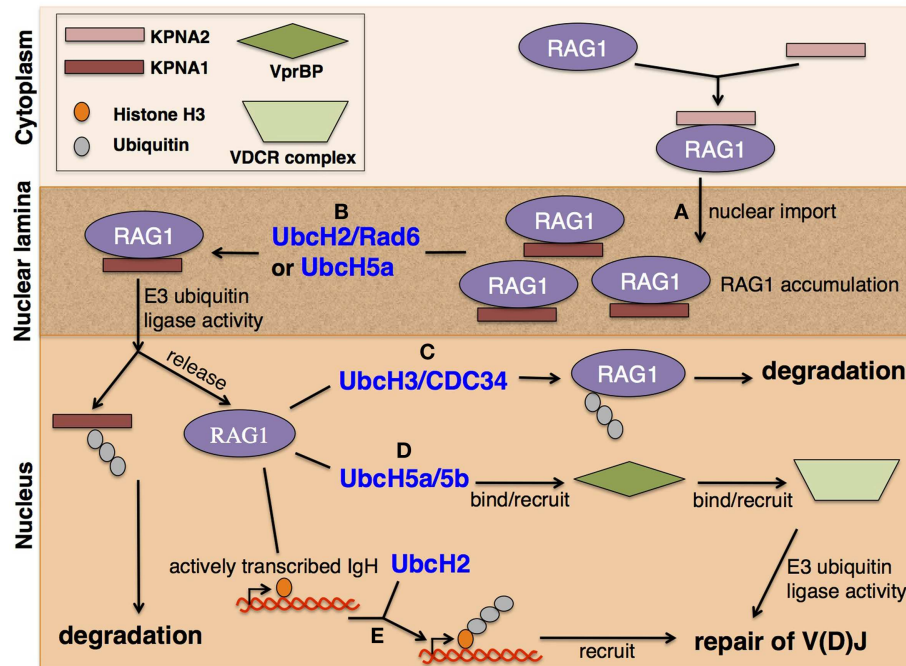
### RAG1 AS AN E3-UBIQUITIN LIGASE

RAG1 is the known catalytic component of the RAG complex, responsible for DNA binding and cleavage during V(D)J recombination [reviewed in Ref. (9)]. Core RAG1 is defined to include all necessary regions required for V(D)J recombination activity. Briefly, a well-defined nonamer-binding domain binds the nonamer sequence of the RSS, as the name implies. A central region (amino acids 528–760) includes a heptamer-binding domain and RAG2 interacting region, which is thought to involve the zinc finger region B (ZnB). Finally, three amino acids (D600, D708, and E962), known as the DDE motif, are important for DNA cleavage. Recently, interest in the significance of the non-core regions of RAG1 has increased, especially with the observation of the conserved N-terminal residues. Utilizing extrachromosomal recombination substrates and deletion and mutation analysis, the need for non-core RAG1 to enhance V(D)J recombination and fidelity has been suggested (10). The ZnA region of RAG1 includes an N-terminal RING domain that acts as an E3-ubiquitin ligase, with the potential to ubiquitinate a panel of targets for various downstream events (11–17), as well as the area which interacts with histone 3 (see below). The ZnA region is able to homodimerize, which may be significant for the E3-ubiquitin ligase activity of RAG1 and/or its regulation (14).

Through an *in vitro* experiment, the N-terminal RING domain of RAG1 has been shown to be capable of and necessary for mono-ubiquitinating test substrate S-protein in the presence of E2 enzymes UbcH10 or UbcH4 (11–17). Point mutations within the RING region, as well as deletion of half of the RING domain, strongly reduce the ubiquitination activity of the wild-type (WT) construct. Poly-ubiquitination was also observed in the presence of E2 UbcH5b, but in the absence of S-protein, suggesting that RAG1 has an auto-ubiquitination capacity. Because the RING domain of RAG1 is dispensable for V(D)J recombination activity, it is possible to conclude that RAG1 has an alternate enzymatic activity, though likely indirectly involved in V(D)J recombination. As noted above, RAG1 was identified as a potential E3-ubiquitin ligase utilizing a synthetic assay. Subsequent studies examining the RAG1 RING domain both *in vitro* and *in vivo*, however, were able to identify specific target substrates ubiquitinated by RAG1, and the nature of these ubiquitinations (11–16). It appears that the RAG1 RING domain is capable of interacting with multiple E2 ubiquitin-conjugating enzymes, with the specific E2 interaction possibly determining the subsequent ubiquitinated substrate(s). Accordingly, possible regulatory mechanisms for RAG1 and RAG2 activity and protein stabilization, as well as a range of alternative downstream pathways, have tentatively been identified.

Ubiquitinated proteins are most often recognized and degraded by the 26S proteasome. However, RAG1's unique RING domain structure, coordinating three Zn ions as opposed to the standard two of canonical RING domains, along with its ability to interact with a panel of E2 enzymes, suggest functions either alternate to or in addition to protein degradation (14). These alternative functions possibly are indirectly involved in V(D)J recombination and/or RAG regulation. RAG1 can be ubiquitinated in intact cells for subsequent degradation by the 26S proteasome, as evidenced by ubiquitinated species only being observed in a state of proteasome inhibition (11). Additionally, through an *in vitro* ubiquitination and pull-down assay, the RAG1 N-terminal region, containing the RING domain, Zn-binding domain, and basic upstream region, was observed to undergo auto-ubiquitination, specifically in the presence of E2 UbcH3/CDC34 (Figure 1C). In an experiment utilizing CH<sub>3</sub>-ubiquitin, which is unable to form poly-ubiquitin chains, it was shown that auto-ubiquitination of RAG1 primarily occurs at one conserved lysine residue, K233. Poly-ubiquitination was observed with WT ubiquitin, but does not necessarily depend on K48-linkage as poly-ubiquitin chains were seen with a K48R ubiquitin mutant. Finally, core RAG1 is more active than full-length protein *in vitro*, suggesting the importance of RAG1 auto-ubiquitination to potentially modulate its own turnover (11). With the RING domain and its E3-ubiquitin ligase activity missing, auto-ubiquitination of RAG1 cannot occur which may lead to the observed heightened activity of core RAG1, and possible promiscuous over-activity.

Beyond auto-ubiquitination and subsequent degradation, the RAG1 RING domain can interact with other E2 enzymes to ubiquitinate substrates involved in V(D)J recombination. One such substrate is histone H3 (13, 15). It was shown that the N-terminal region of RAG1 directly interacts with histone H3 and an intact RING domain is required for mono-ubiquitination of histone H3 *in vitro* and *in vivo* (13). Through experiments using both mutated and truncated proteins, it was determined that the N-terminal domains of both RAG1 and endogenous histone H3 directly interact. Structural analysis reveals RING mutations do not affect overall protein folding, just RAG1 activity. Furthermore, extrachromosomal V(D)J recombination assays demonstrate that point mutations within the RING domain of RAG1 cause deficient DNA joining, but not cleavage, at the endogenous IgH locus *in vivo*. It is interesting to note that patients with Omenn syndrome, a condition that results in combined immune deficiencies including reduced efficiency of V(D)J recombination, also harbor point mutations in the RAG1 RING domain (18, 19). Though inefficient V(D)J recombination from a RING domain mutation may be a readout of disrupted regulation from RAG1 auto-ubiquitination, ubiquitination of histone H3 may be an important factor during V(D)J recombination to (a) tag the RSS breaks to recruit DNA repair proteins, (b) destabilize the nucleosome or remodel the chromatin for DNA accessibility for repair, (c) promote RAG complex eviction to allow joining machinery to complete coding and signal joints, or (d) promote cell-cycle arrest (13). These possibilities are not mutually exclusive, and all have the potential to provide alternative pathways to mediate effective V(D)J recombination. Rather than one exclusive mechanism, it is likely a



**FIGURE 1 | RAG1 localization for V(D)J recombination.** (A) RAG1 is imported to the nucleus through nuclear pores upon KPNA2 binding to RAG1 NLS signal, located in the core region. RAG1 is thought to be retained within the nuclear lamina by binding KPNA1 upstream of the RING domain. RAG1 accumulation in the nuclear lamina promotes RAG1-dependent ubiquitination of KPNA1 in the presence of (B) E2 UbH2/Rad6 or UbH5a, allowing simultaneous degradation of KPNA1 by the 26S proteasome and release of RAG1 into the nucleus. RAG1 is capable of E3-ubiquitin ligase activity through its RING domain in

the presence of a panel of E2 ubiquitin-conjugating enzymes. (C) Whereas E2 UbH3/CDC34 promotes auto-ubiquitination of RAG1 for protein degradation, interaction with (D) E2 UbH5a/5b enables RAG1 to bind with VprBP, which then binds to the VDCR complex to recruit its E3 activity, possibly to repair DNA breaks caused during V(D)J recombination. (E) E2 UbH2 allows RAG1 to recognize actively transcribed DNA that was subject to DNA cleavage by core RAG1 by binding and ubiquitinating histone H3, specifically acetyl-H3.3 S31p, thereby tagging the DNA region for repair by NHEJ factors.

combination of the proposed models that is responsible for proper V(D)J recombination.

Subsequent investigations describe histone variant H3.3 as the target of RAG1 RING ubiquitination (15). Mutant RAG1 protein experiments and *in vitro* assays indicate binding between the RAG1 RING domain and the N-terminus of histone H3.3. From mass spectrometric analysis of H3 modifications, it is thought that acetylation and phosphorylation of H3.3 (acetyl-H3.3 S31p) are a mark of active chromosomes in mitotic cells, activate the histone as a substrate for RAG1 RING-dependent ubiquitination, and are upregulated during V(D)J recombination. Accordingly, acetyl-H3.3 S31p could act as a tag for recombining loci catalyzed by core RAG1 during V(D)J recombination (Figure 1E). Additionally, H3.3 was shown to have multiple sites of mono-ubiquitination in the presence of E2 UbH2. It is possible that after DNA cleavage during V(D)J recombination, the RAG1 RING domain binds acetyl-H3.3 S31p to target the break site and recruit DNA repair complexes to complete the joining step and ensure faithful V(D)J recombination. RAG1 was also identified as an E3-ubiquitin ligase complex member, supporting the potential role of ubiquitinated histones as marking breaks to recruit repair proteins (12). Evidence for this role includes the observation that full length, but not core, RAG1 with core or full-length RAG2 co-purifies with a

complex containing VprBP, DDB1, Cul4A, and Roc1 (VDCR complex) *in vitro*, which is known to act as a recruitment scaffold for repair proteins (Figure 1E).

Although there are multiple reports of RAG1 being capable of E3-ubiquitin ligase activity, there are discrepancies as to with which E2 enzyme the RAG1 RING domain interacts. Since the E3 enzyme of most ubiquitin reactions largely determines the specific substrate to be ubiquitinated, it is interesting to note that it is the E2 enzyme instead that appears to target the downstream ubiquitinated product in the context of E3 RAG1 (Figure 1). In previously described studies, the presence of different E2 enzymes results in different substrates of RAG1 RING. This observation could be a consequence of differing experimental designs and the lack of a reliable *in vivo* system. Alternatively, the results could properly depict the broad range of potential proteins subject to ubiquitination by the RAG1 RING E3 ligase activity for purposes other than protein degradation.

As a potential model, upon RAG1 nuclear import (Figure 1A), the interaction and ubiquitination of KPNA1 releases RAG1 from the nuclear lamina with the assistance of E2 UbH2/Rad6 or UbH5a (20) (Figure 1B). The freed nuclear RAG1 is then able to auto-ubiquitinate in the presence of E2 UbH3/CDC34, regulating its own protein levels by marking itself for degradation (11)

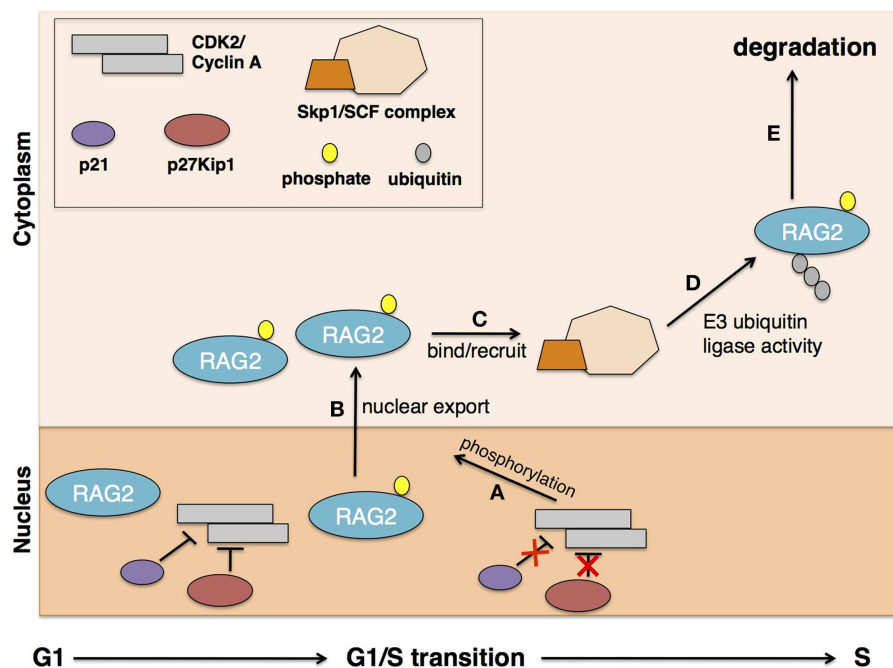
(Figure 1C). Alternatively, RAG1 can act with E2 UbcH2 to target and ubiquitinate histone H3 of actively transcribing DNA for both processing by V(D)J recombination and for tagging the cleaved DNA to recruit proteins involved in DNA repair to promote robust and efficient NHEJ (13) (Figure 1E). Simultaneously, RAG1 with E2 UbcH5a/5b can interact with VprBP of the VDCR complex, an E3-ubiquitin ligase complex, to recruit repair proteins during V(D)J recombination (12) (Figure 1D). The proposed model presents another level for the cell to regulate RAG1 by regulating E2 enzymes available for use. For example, if E2 UbcH2/Rad6 and UbcH5a are not available during S phase of the cell cycle, RAG1 would not be able to ubiquitinate KPNA1 and therefore unable to release itself from the nuclear lamina to perform its catalytic role. Regulation of these E2 enzymes prevents promiscuous RAG1 activity outside of appropriate V(D)J recombination events. Similarly, the presence of E2 UbcH3/CDC34 or UbcH2 are important in regulating RAG1 protein stabilization or recruiting repair proteins for efficient V(D)J recombination, respectively. The second activity of RAG1 as an E3-ubiquitin ligase, in addition to catalyzing DNA breaks, provides alternative mechanisms in modulating V(D)J recombination that include multiple pathways for RAG1 activity regulation.

#### REGULATION OF RAG2 BY UBIQUITINATION

Though the function remains rather elusive, RAG2, like RAG1, has a defined core region (amino acids 1–383) located at the

N-terminal end and with core RAG1, is required for V(D)J recombination [reviewed in Ref. (17)]. The RAG2 core is known to be critical for DNA cleavage and enhances DNA binding and specificity of the RAG complex, probably through its interaction with RAG1, as RAG2 itself has little or no DNA-binding activity. The RAG2 non-core region includes a PHD domain that specifically binds H3K4me3 to guide RAG2 to active chromatin and enhances the catalytic activity of the RAG complex. An important characteristic of RAG2 is its periodic accumulation and degradation in relation to the cell cycle (21). RAG2 is phosphorylated at residue Thr-490, located in the C-terminal non-core region, by the cell-cycle-dependent protein cyclin A/CDK2. Cyclin A/CDK2 is upregulated at the G1/S phase transition and maintained through entry into M phase. It is known that p27Kip1, a cyclin-dependent kinase inhibitor, blocks the activity of cyclin A/CDK2, in effect stabilizing nuclear RAG2 protein levels (21, 22). It is therefore hypothesized that at the G1/S phase transition, p27Kip1 is degraded to allow phosphorylation of RAG2 at Thr-490 by cyclin A/CDK2, leading to RAG2 protein degradation and thereby regulating RAG2 protein levels in a cell-cycle-dependent manner (Figure 2A).

Specifically, it was shown that RAG2 degradation is mediated by ubiquitination and subsequent activity of the 26S proteasome, as poly-ubiquitinated RAG2 species were observed upon treatment with a specific 26S proteasome inhibitor (23). Furthermore, ubiquitination assays comparing full length, T490A mutant, and C-terminal-deleted RAG2, with and without 26S proteasome



**FIGURE 2 | Cell-cycle-dependent RAG2 localization for V(D)J**

**recombination.** Periodic RAG2 accumulation is dependent on cyclin A/CDK2 activity. In the G1 cell-cycle phase, cyclin A/CDK2 is inhibited by p27Kip1 and p21, and RAG2 protein is stable and abundant in the nucleus. Upon transition into S phase, p27Kip1 and p21 are degraded, allowing (A) cyclin A/CDK2 activity to phosphorylate RAG2 at Thr-490 residue. This phosphorylation event

triggers (B) nuclear export of RAG2 where it can (C) bind Skp2 in the cytoplasm to recruit the SCF complex. The SCF complex is a known E3-ubiquitin ligase complex that (D) ubiquitinates RAG2 for (E) rapid degradation by the 26S proteasome. This model explains the observed periodic accumulation of RAG2 during G1, coinciding proper cell-cycle activity for efficient V(D)J recombination.



inhibition, suggest that sites for ubiquitin ligase interaction and ubiquitination are located at the N-terminal region of RAG2. Though the C-terminal region was observed to be inhibitory to ubiquitination and degradation of RAG2, phosphorylation of the Thr-490 residue abrogates this C-terminal inhibitory activity. These findings demonstrate self-regulatory functions of RAG2 through interactions of different regions of the protein. As alluded to earlier, it was observed that RAG2 localizes in both the cytoplasm and the nucleus, and this subcellular localization is cell-cycle dependent. Western blot analysis shows cytoplasmic RAG2 is less stable than nuclear RAG2 (23). Along with active p27Kip1, the C-terminal region helps to retain RAG2 within the nucleus. However, post-translational phosphorylation of Thr-490 of RAG2 seems to abrogate the C-terminal inhibitory function, allowing nuclear RAG2 export into the cytoplasm (**Figure 2B**).

Cytoplasmic localization renders RAG2 accessible for ubiquitination and degradation. It was observed that phosphorylated Thr-490 of RAG2 is an interacting site for another cell-cycle regulator, Skp2-SCF (24). Via biochemical fractionation, E2 Cdc34, commonly associated with the SCF family of E3-ubiquitin ligases, was identified as stimulating RAG2 ubiquitination. Further biochemical assays determined Skp2-SCF as the specific E3-ubiquitin ligase complex for RAG2 ubiquitination, with confirmation from subsequent *in vivo* mutation and knock-down experiments. G1 and S/G2/M phase cells collected from Skp2-deficient mice demonstrated that the periodic accumulation of RAG2 is dependent on Skp2, as RAG2 protein levels from S/G2/M phase were elevated in Skp2-deficient mice. Other *in vitro* experiments confirmed the dependence on the Skp2-SCF complex in regulating RAG2 protein with the cell cycle. Skp2 was also shown to bind Skp1 of the SCF complex, which recruits the complex's E3 activity to RAG2 for ubiquitination and degradation. The SCF complex possesses alternative roles for targeting other cell-cycle factors such as p27Kip1 and p21 for degradation (**Figures 2B,C**). Interestingly, both p27Kip1 and p21 are inhibitors of cyclin A/CDK2 and are therefore negative regulators of RAG2 degradation. The importance of SCF E3-ubiquitin ligase activity is highlighted by the range of target proteins, which all ultimately converge to regulate the activity of RAG2 and permit proper V(D)J recombination.

The findings from experiments investigating RAG2 activity and degradation have emphasized RAG2's role in regulating V(D)J recombination activity to appropriate phases of the cell cycle. DNA breaks during S phase are potentially harmful for cells as they can lead to translocations and lymphomas when misrepaired by homologous recombination (HR). It is therefore crucial to limit V(D)J recombination activity within the G1 phase; the restriction appears to be controlled by RAG2 localization and stabilization. RAG2 nuclear accumulation is only observed during G1, presumably because upon transition to S phase, p27Kip1 is degraded, relieving suppression of cyclin A/CDK2 activity and permitting phosphorylation of RAG2 at Thr-490 (21, 22). The phosphorylation at Thr-490 allows for nuclear export where the phosphorylated residue interacts with Skp2 to recruit the E3-ubiquitin ligase activity of the SCF complex (24). Ubiquitination of RAG2 allows for rapid degradation of the protein upon entering S phase, thereby halting any potential off-target activities of RAG (**Figure 2**).

The periodic accumulation and degradation of RAG2 has significant implications for V(D)J recombination activity. Since RAG2 together with RAG1 is required for catalysis of V(D)J recombination events, its nuclear retention during G1 is important. However, upon S phase entry, RAG2 is rapidly degraded to prevent non-specific DNA cleavage events. Without RAG2, RAG1 can remain bound to DNA where its RING E3-ubiquitin ligase activity can ubiquitinate histone H3 and/or VprBP, as discussed earlier. Though RAG1 is catalytically compromised in the absence of RAG2 for V(D)J recombination activity, it still retains its E3-ubiquitin ligase activity. This proves significant as ubiquitination of H3.3 possibly tags DNA for repair while ubiquitination of VprBP is thought to recruit the E3 activity of the VDCR complex. Together, degradation of RAG2 during the G1/S phase transition simultaneously halts V(D)J recombinase activity and recruits NHEJ repair proteins to sites of DNA breaks for efficient and productive V(D)J recombination.

### AID-MEDIATED CLASS SWITCH RECOMBINATION AND SOMATIC HYPERMUTATION

Upon completion of V(D)J recombination in the bone marrow, immature IgM<sup>+</sup> B cells migrate to secondary lymphoid tissues for further DNA alterations events, namely CSR and SHM, which are mediated by AID. Following antigen-dependent activation, germinal center B-lymphocytes undergo CSR to generate antibodies with different effector functions, and SHM to increase the affinity of the antibody for its cognate epitopes, a process also known as affinity maturation (6, 25, 26). The IgH locus has multiple constant region genes (C<sub>H</sub>) that are preceded by G-rich switch sequences (S) and subject to CSR. In mice, there are eight sets of C<sub>H</sub> exons organized as 5'-V(D)J-C<sub>μ</sub>-C<sub>δ</sub>-C<sub>γ</sub>3-C<sub>γ</sub>1-C<sub>γ</sub>2b-C<sub>γ</sub>2a-C<sub>ε</sub>-C<sub>α</sub>-3', and each S sequence has its own cognate promoter that is influenced by enhancer elements, such as the 3' regulatory region enhancer (3' RR) (5). This allows for transcription of stimulated B cells at S<sub>μ</sub> and another S sequence, such as S<sub>γ</sub>1, S<sub>γ</sub>3, S<sub>ε</sub>, or S<sub>α</sub> (C<sub>δ</sub> lacks an upstream switch sequence), resulting in selection of a constant region that now encodes IgG1, IgG3, IgE, or IgA, respectively. Transcription-dependent generation of a DNA DSB mediated by AID at the donor S<sub>μ</sub> sequence and downstream acceptor switch sequence leads to a recombination-deletion event that removes the intermediate sequences and propagates both S<sub>donor</sub>-S<sub>acceptor</sub> synapsis and catalysis of CSR [reviewed in Ref. (1, 4)]. Additionally, B cells carry out SHM by introducing mutations, also catalyzed by AID, in the variable regions of the IgH and IgL loci, which, when synthesized, are in physical contact with the antigen during an immune challenge. The mutations in these segments occur at a frequency much higher than in other regions of the genome. The mutations are initiated at RGYW motifs by AID and spread as a consequence of downstream events orchestrated by the mismatch repair (MMR) and base excision repair (BER) pathways (27, 28) [reviewed in Ref. (6, 29, 30)]. Understanding the molecular mechanism by which transcription within various regions of the Ig locus coordinates with the mutagenic activity of AID to generate and regulate programmed DNA lesions has been a challenge. Comprehension is, however, necessary to fully understand AID and its tumorigenic potential upon misregulation.

## ACTIVATION-INDUCED CYTIDINE DEAMINASE AND PROTEIN (IN)STABILITY THROUGH UBIQUITINATION

Activation-induced cytidine deaminase is a single-strand cytidine deaminase that utilizes transcription-dependent mechanisms to generate single-strand DNA (ssDNA) structures that allow mutagenesis of target DNA substrates of B cells (31). Chromatin immunoprecipitation (ChIP) of AID from CSR-stimulated B cells followed by high throughput sequencing of AID-associated DNA fragments (ChIP-seq) reveals that AID can bind various regions of the B-cell genome, inside and outside the Ig loci, presenting opportunities for genomic instability (32). Therefore, understanding the physiological pattern and distribution of AID-generated mutations at AID's target sequences is vital. Changes in AID mutation distributions at S regions can decrease CSR efficiency and generate DSB intermediates causing oncogenic IgH translocations while similar aberrant mutation patterns at the variable regions during SHM can alter antibody specificity for antigen (6, 33–36). Indeed, post-translation modifications and co-factors of AID have been identified and proposed to affect its activity through multiple mechanisms, such as (a) stimulation of AID's DNA deamination activity, (b) linking AID to the Ig transcription machinery, (c) establishing a physiological DNA deamination pattern on both strands of DNA, and (d) linking AID to the downstream DNA repair machinery [reviewed in Ref. (4)]. Two separate studies investigated the regulatory role of post-translation AID ubiquitination during CSR and SHM (37, 38). Unlike RAG2, AID protein stability is not associated with phases of the cell cycle, but rather with subcellular localization. Utilizing expression constructs, AID mutants, and pulse-chase experiments, the Reynaud group showed that in mouse B cells and 293T cells, nuclear AID is subject to rapid turnover upon poly-ubiquitination (37). Nuclear-restricted AID was shown to have enhanced mutagenic activity in both Ig and non-Ig loci, demonstrating the importance of controlling AID off-target activity. As no specific lysine residue was determined to be the target of ubiquitination, it remains unclear whether several sites are poly-ubiquitinated or if N-terminal residues of AID are targeted. This work highlights the multilayer regulation of AID function, including protein stability and turnover (37).

Since AID is known to promote oncogenic mutations in the B-cell genome (4, 39, 40) and its protein levels are likely controlled by ubiquitin-mediated degradation (37), it is crucial to identify mechanisms of AID ubiquitination, specifically in the context of the transcription complex. To this end, the Papavasiliou group recently discovered that the RING E3 ligase RNF126, with E2 UbcH5b, can mono-ubiquitinate AID in cell-free assay conditions and in 293T cells (38). Though the functional implications of RNF126-mediated AID ubiquitination were not investigated, the authors provide several mechanistic insights concerning the potential role of AID ubiquitination. Auto-ubiquitination and ubiquitin binding, due to the presence of an N-terminal ubiquitin-binding domain of RNF126, may prevent recruitment of PCNA and translesion polymerase, which inhibits spreading of AID-generated mutations. Alternatively, RNF126 could be involved in regulating transcription initiation at promoters of AID target genes, presenting important implications for AID targeting, discussed below. Better understanding of RNF126 in AID regulation

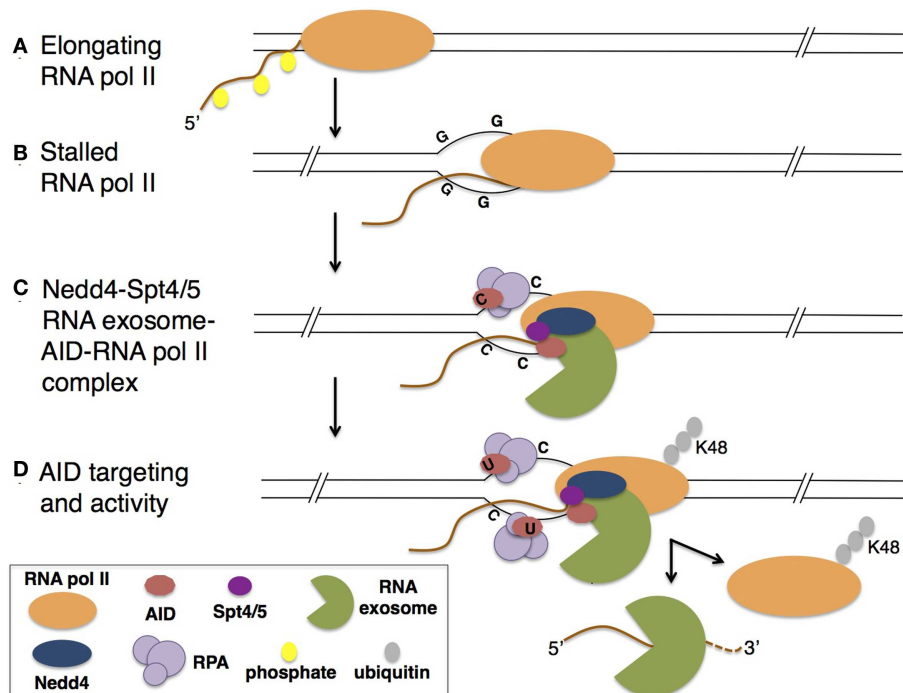
will be dependent upon generation of RNF126-deficient mouse models (38).

## UBIQUITINATION OF AID-ASSOCIATED RNA POLYMERASE II

Beyond post-translation modifications to regulate protein stability, co-factors are also important for proper AID activity. AID is proposed to bind the paused and/or stalled state of RNA polymerase II (RNA pol II), consistent with its transcription-dependent activity (4, 39, 40). Whereas “paused RNA pol II” is bound to DNA and is positioned at DNA sequences proximal to genic transcription start sites (TSSs) prior to entering transcriptional elongation (41), “stalled RNA pol II” is positioned on template DNA during RNA pol II elongation (42). RNA exosome and Spt5 are co-factors of both RNA pol II and AID in B cells (43, 44). Though the 3' → 5' RNA exonuclease complex RNA exosome is predominantly associated with the stalled RNA pol II complex, Spt5 is associated with both paused and stalled RNA pol II complexes (42). ChIP-seq data show AID-bound sequences have high occupancy by RNA pol II molecules that are either in the elongation phase or in the paused or stalled state (43). Consistent with these observations, RNA pol II is enriched at various regions of IgS sequences, an AID target (45). The switch sequences are enriched with AID and Spt5 (32, 43). These results present a role for the transcription machinery in AID targeting and the importance of regulating the transcription complex to prevent promiscuous activity of AID. Below, we will discuss the role of ubiquitination in regulating these processes.

Many AID-induced mutations occurring during CSR (36) or SHM (46) are significantly downstream from the transcription start sites of V genes or switch sequences (as opposed to 100 bps downstream of TSS where RNA pol II promoter-proximal pausing occurs). This downstream localization is also true for other AID target genes (47). Thus, it is likely that elongation-stalled RNA pol II is an adequate complex for recruitment of AID during SHM and CSR. RNA exosome recruitment to stalled RNA pol II requires the presence of a nascent RNA with a free 3'-end. Stalled RNA pol II complexes occur during encounter with various obstacles caused by (a) the presence of antisense transcription, (b) secondary DNA structures including those caused by G-richness on the non-template strand of S sequences, (c) variation in the levels of elongation promoting chromatin modifications, and (d) the presence of mutations or DNA lesions on the template DNA (4, 39, 40). Resolution of the stalled complex includes mechanisms of backtracking or early termination of the complex, which dissociates the RNA from RNA pol II (41), or ubiquitin-mediated destabilization of RNA pol II (42, 48). These mechanisms reveal a free 3'-end RNA substrate for RNA exosome, a co-factor of AID.

Recently, it was shown that E3-ubiquitin ligase Nedd4 identifies and ubiquitinates AID-associated RNA pol II (49). Through immunoprecipitation experiments, AID was shown to interact with RNA pol II, stalling factor Spt5, and RNA exosome. AID-associated RNA pol II is poly-ubiquitinated in a Nedd4-dependent fashion. B cells obtained from Nedd4-mutant mice demonstrate increased stability of IgH germline transcripts that are expressed from DNA switch sequences subject to AID mutations, indicating that Nedd4 activity promotes processing of germline transcripts, possibly by recruiting AID. Finally, B cells from Nedd4-mutant



**FIGURE 3 | Model for E3 ligase Nedd4 in mediating CSR. (A)** During transcription of the B-cell genome, RNA pol II in the elongation complex encounters **(B)** a transcriptional block, such as secondary structures formed in the G-rich switch regions of Ig genes. **(C)** Stalled RNA pol II exists in a complex with interacting E3 ligase Nedd4, Spt4/5, RNA exosome, and AID.

RPA-associated AID targets the non-template DNA strand.

**(D)** K48-poly-ubiquitinated RNA pol II is released, relieving the 3' end of the RNA transcript available for RNA exosome degradation. The release of the DNA/RNA hybrid allows AID to target both template and non-template ssDNA (73).

mice are impaired in CSR and have a compromised mutation rate in the 5' S $\mu$  sequence, an AID target. These observations highlight the requirement of Nedd4 and its ubiquitin ligase activity for proper AID-associated RNA pol II activity during antibody diversification mechanisms (Figure 3).

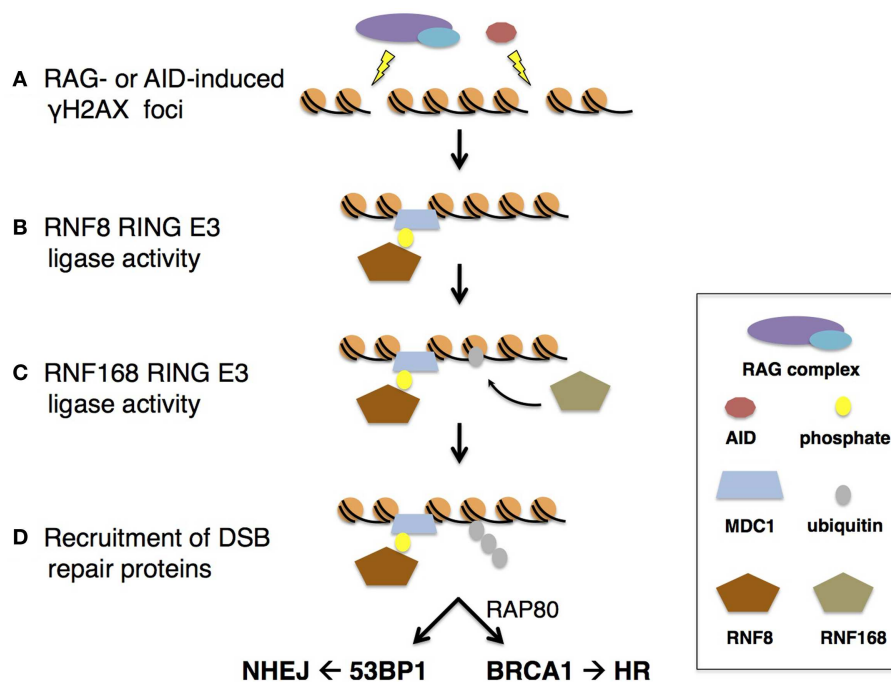
The requirement of AID-associated RNA pol II ubiquitination has implications for AID targeting. First, ubiquitin-mediated RNA pol II destabilization exposes the 3' end of the RNA pol II-associated nascent transcript for degradation by RNA exosome. Nedd4 activity therefore could contribute toward the generation of ssDNA substrates on both the template and non-template strands of DNA for AID deamination by triggering RNA exosome activity (Figure 3), possibly as an alternative mechanism to RNA pol II backtracking. In addition to facilitating targeting of AID activity to specific regions of the genome, Nedd4 may also prevent accumulation of stalled RNA pol II at AID target regions in the B-cell genome to prevent generation of aberrant DNA DSBs. Collision of the stalled RNA pol II with replication machinery may be responsible for DNA DSBs and chromosomal translocations. Whether Nedd4 ubiquitination activity is required for RNA pol II destabilization during SHM to promote mutations in variable region genes or at oncogenic targets of AID is important to determine.

In addition to direct ubiquitination of AID to regulate its nuclear localization and stability, ubiquitination also appears to indirectly control AID targeting to its physiological DNA substrates by facilitating AID-associated RNA pol II destabilization

to reveal substrates for RNA exosome, an AID co-factor. Protein stability and specific targeting of nuclear AID are vital for effective CSR and SHM, and therefore regulation of both processes is required to prevent off-target activity and genomic instability.

## REPAIR OF DNA DOUBLE-STRAND BREAKS DURING B-CELL DEVELOPMENT

As immature B cells are subject to programmed DNA DSBs during V(D)J recombination, CSR, and SHM during development, appropriate repair of these DSBs is essential for proper B-cell maturation and prevention of disease and oncogenesis. Importantly, both RAG- and AID-induced DSBs depend on NHEJ, not HR, for repair (1, 17). DSBs are initially recognized by the Ku70/Ku80 heterodimer to recruit kinase DNA-PK [reviewed in Ref. (50)]. Other DSB repair factors are then recruited, including XRCC4-ligase IV, Pol  $\mu$ , Pol  $\lambda$ , and TdT in lymphocytes. Phosphorylated histone H2AX on Ser139 ( $\gamma$  H2AX) by DNA-PK is a well-known mark for DNA damage and recruits MDC1 (mediator of DNA damage checkpoint protein 1) via interaction with the BRCT motif (51). Upon interacting with phosphorylated ATM through its FHA domain, MDC1 both stabilizes the histone modification and amplifies the  $\gamma$  H2AX signal (52). ATM then phosphorylates TQXF motifs of MDC1. It is at this point in the NHEJ repair pathway that ubiquitination assumes its prominent role (53, 54). The initial recognition and recruitment of repair proteins to DSBs is heavily dependent on phosphorylation events (e.g., kinase



**FIGURE 4 | RNF8- and RNF168-mediated repair of RAG- or AID-induced DSBs.** (A) DSBs from RAG or AID activity result in  $\gamma$ H2AX foci formation to recruit MDC1. (B) RNF8 recognizes ATM-dependent phosphorylated MDC1 via its FHA domain. The RING domain of RNF8 then ubiquitinates DSB-associated chromatin. (C) RNF168 recognizes

ubiquitin-modified breaks via its MIU motif. RNF168 is recruited to DSB sites and (D) extends the RNF8-mediated ubiquitin conjugates. The RNF8/RNF168-dependent poly-ubiquitinated products are then recognized for repair completion by BRCA1 through RAP80 for HR, or 53BP1 for NHEJ.

activity and autophosphorylation of DNA-PK), but it is the downstream repair events that appear to be dependent on E3-ubiquitin ligases (55).

#### ROLE OF E3 LIGASES RNF8 AND RNF168 DURING REPAIR OF DNA DSBs

At the crossroads between phosphorylation- and ubiquitin-dependent events in DSB repair, phosphorylated MDC1 is recognized by the E3 ligase RNF8 through its FHA (forkhead associated) domain (56). Two separate groups described RNF8 binding to MDC1 for ubiquitinating damaged-associated histones (57, 58). Mailand et al. used a bioinformatics approach to panel motifs of known DSB regulators to identify RNF8. Its function was tested and it was observed to colocalize with  $\gamma$  H2AX at DSBs; colocalization was abrogated in the absence of MDC1 (57). To further characterize the relationship between RNF8 and MDC1, Mailand used a combination of biochemical and real-time imaging techniques to reveal the direct interaction of RNF8 and MDC1, which requires the FHA domain and TQXF motif of RNF8 and MDC1, respectively. Furthermore, RNF8 was shown to rapidly accumulate at DSBs with the same kinetics as MDC1, preceding 53BP1 and BRCA1 accumulation, suggesting that RNF8 functions upstream of 53BP1 and BRCA1, factors involved with NHEJ and HR, respectively. The second group, Huen et al., used a tagged-RNF8 construct to biochemically determine its colocalization with  $\gamma$ H2AX and other known damage response proteins, such as MDC1 and 53BP1, to DSBs (58). Consistent with the previously described study, RNF8 was shown to function downstream of  $\gamma$ H2AX and

MDC1 recruitment and to interact directly with phosphorylated TQXF motifs of MDC1 via its FHA domain. Both groups demonstrate the necessity of both the FHA and RING domains for complete RNF8 function. Importantly, the RING domain is capable of ubiquitin ligase activity on H2A (57) and H2AX (58) in *in vitro* ubiquitination assays. In the presence of E2 Ubc13, RNF8 is capable of mono- and di-ubiquitinating H2AX. Taken together, these findings present evidence for a functional link between RNF8 and H2A, possibly modifying histones to reveal buried substrates necessary for downstream interactions by NHEJ repair proteins, such as 53BP1.

Subsequent to the two previous studies, the Durocher group showed that RNF8 also impairs 53BP1 focus formation (53). Depleting RNF8 abrogated 53BP1 foci without disrupting MDC1 foci. Further investigation revealed that the N-terminal FHA domain and C-terminal RING domain of RNF8 are both necessary for 53BP1 focus formation as mutation in either domain abolishes 53BP1 focus formation. This study highlights the significance of  $\gamma$ H2AX- and MDC1-dependent RNF8 response to DSBs in recruiting 53BP1 accumulation. However, direct interaction between phosphorylated MDC1 and 53BP1 was not observed. This is not unexpected, as 53BP1 is known to recognize H4K20me2, suggesting an additional mediator exists downstream of RNF8 to recruit 53BP1 (59).

Two groups utilizing similar genome-wide siRNA screens identified RNF168 as the additional mediator for 53BP1 focus formation (60, 61). Both groups confirmed RNF168 acts downstream

of RNF8 in the repair pathway, as depletion of MDC1 and RNF8 abrogated RNF168 foci, but the foci were unaffected by depletion of 53BP1 or BRCA1 (61). Furthermore, knock-down of RNF168 results in RNF8 accumulation at DNA damage sites but RNF8-dependent ubiquitinated chromatin is unstable, suggesting transient activity of RNF8 that is stabilized by RNF168 (60). As 53BP1 foci were dependent on both the N-terminal RING and two MIU (motif interacting with ubiquitin) domains of RNF168, the RING domain of RNF168 was then tested for ubiquitin ligase activity. In *in vitro* ubiquitin assays, RNF168 was indeed capable of E3 activity, with direct interaction with Ubc13, the only known E2 capable of catalyzing formation of K63-ubiquitin chains. Subsequent experiments confirmed RNF168 specifically ubiquitinates H2A type histones, including H2A and H2AX, with E2 Ubc13 to form K63-ubiquitin chains. The poly-K63-ubiquitinated histones were shown to be dependent on both RNF8 and RNF168. These findings suggest a model whereby RNF8 is first recruited to sites of DSBs via its FHA domain, recognizing phosphorylated TQXF motifs of MDC1 (**Figure 4B**). RNF8 is then able to initiate the ubiquitination of  $\gamma$ H2AX to recruit RNF168 via its MIU domains to promote K63-ubiquitin chain extension (60) (**Figure 4C**). In this way RNF168 stabilizes and amplifies the transient ubiquitin conjugate signals from RNF8 for recruitment of downstream repair proteins (**Figure 4D**).

Though the sequential recruitment of RNF8 followed by RNF168 is thought to mimic their sequential activity at DSBs for repair (**Figure 4**), recent observations by Mattioli et al. challenge this model. *In vitro* assays confirm the ability of RNF8 and RNF168 to ubiquitinate free histone variants (62). However, in the presence of purified nucleosomes, where the octamer surrounds DNA, only RNF168 is capable of ubiquitinating H2A. Upon further investigation, results suggest a mechanism whereby RNF168 recruitment to sites of DSBs is RNF8-dependent, but RNF8-mediated ubiquitin chain extension is dependent upon mono-ubiquitination by RNF168. If this revised model holds, it will be interesting to determine how RNF168 is initially recruited. Panier et al. suggest RNF168 recruitment involves two “waves” of RNF168 accumulation at DSBs, and describe the role of p97 in removing RNF8-mediated ubiquitin conjugates to unmask H4K20me2 for 53BP1 interaction and recruitment (63).

#### RNF8- AND RNF168-DEPENDENT RECRUITMENT OF 53BP1 TO DSBs

Though the RNF8- and RNF168-dependent recruitment of repair proteins to DSBs is well-documented, the mechanism by which 53BP1 recognizes its substrate, H4K20me2, remains unclear. As mentioned above, RNF168 was shown to preferentially produce K63-ubiquitin chains via its interaction with E2 Ubc13. How, then, is 53BP1 able to recognize RNF8- and RNF168-modified DSBs? Because of its effect on DNA repair and its important role in various ubiquitin-dependent processes in the cell, the ubiquitin-selective segregase p97 was investigated in relation to RNF8 and RNF168 (64, 65). The accumulation of p97, and its adaptor protein NPL4, at DSB sites was shown to be RNF8-dependent but RNF168-independent (66). Moreover, RNF8 RING domain and free nuclear ubiquitin must be present for p97 foci formation. *In vivo* co-affinity purification assays show p97 and RNF8 interact in a complex, with RNF8 directly binding

ubiquitin and p97 interacting with ubiquitinated moieties (67). Real-time recruitment kinetics show p97 accumulation occurring after MDC1 but before 53BP1, supporting observations of MDC1- and RNF8-dependent recruitment for efficient downstream 53BP1 recruitment (66, 67). BRCA1 focus formation was not affected by p97 depletion, confirming the specificity of p97 for 53BP1 recruitment (66).

Though 53BP1 recruitment is mediated by RNF168 ubiquitin ligase activity, 53BP1 interacts with H4K20me2, not ubiquitinated moieties (59). p97 was also observed to bind to and regulate the turnover of RNF8-dependent K48-ubiquitin chains (67). This suggests a mechanism by which RNF8 catalyzes K48-ubiquitin chains to recruit p97 through its ubiquitin adaptor UFD1–NPL4. The K48-ubiquitin conjugates are then removed by the ATPase-driven segregation activity of p97, which causes a rearrangement of the DSB-associated chromatin complex, possibly revealing buried H4K20me2, allowing interaction with and recruitment of 53BP1 (67). This provides insight into the factors directly mediating 53BP1 focus formation downstream of RNF8 activity for proper repair by NHEJ. Importantly, the polycomb complex also interacts with H4K20me2 and with higher affinity than 53BP1 (61). However, upon DNA damage and subsequent  $\gamma$ H2AX focus formation, L3MBTL1 (a polycomb complex protein) is poly-K48-ubiquitinated with subsequent increased p97 interaction and decreased H4K20me2 association (66). Taken together, it is possible that RNF8 poly-ubiquitinates the H4K20me2-associated polycomb complex at DSB-associated chromatin to recruit p97–UFD1–NPL4 through direct binding of ubiquitin moieties. The ATPase-driven segregation activity of p97 then mediates the turnover of K48-ubiquitinated polycomb complex, relieving H4K20me2 sites for recruitment of 53BP1. In this way, p97 is essential for connecting the E3-ubiquitin ligase activities of RNF8 to 53BP1 recruitment for NHEJ by revealing H4K20me2 binding sites for 53BP1.

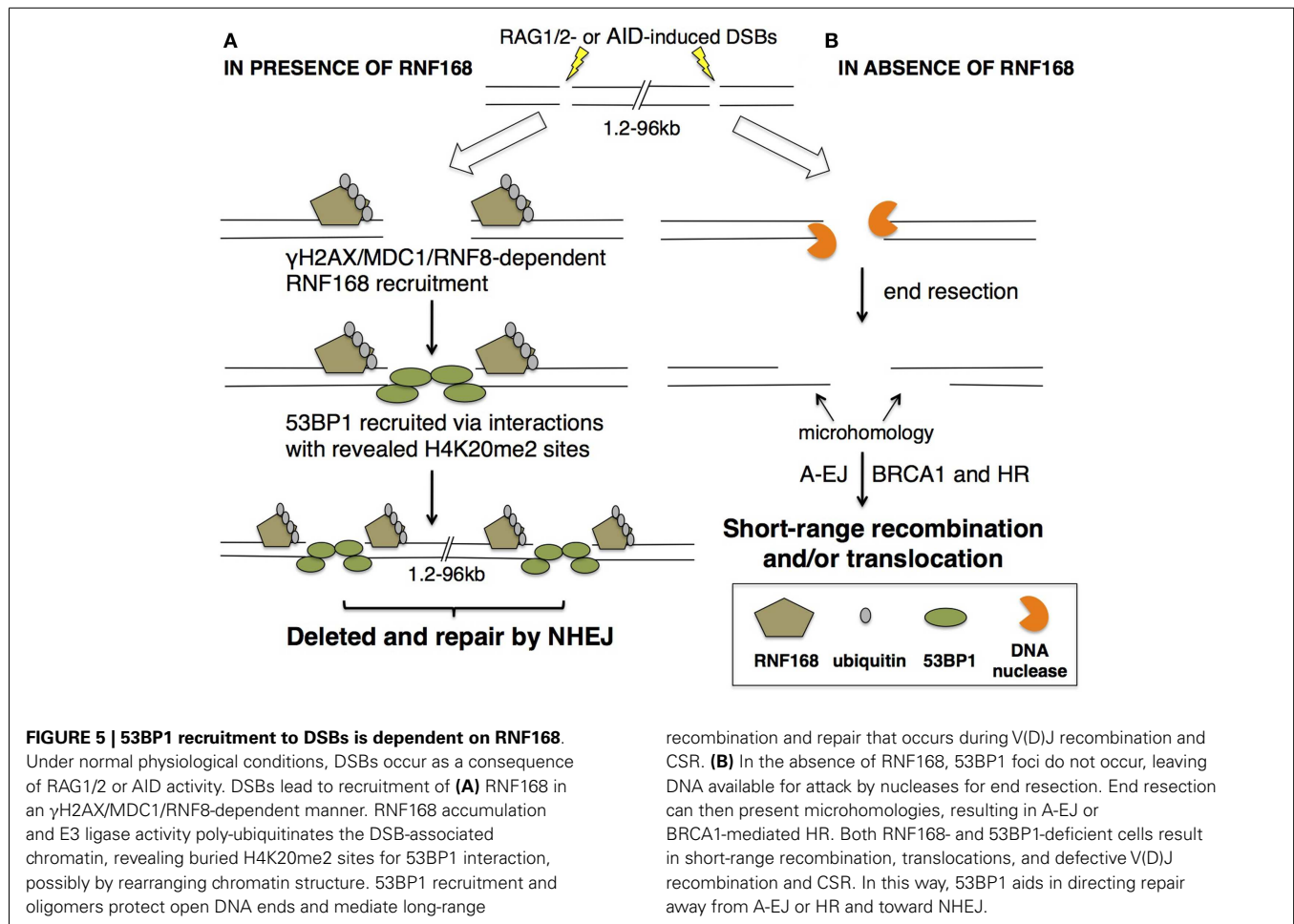
#### 53BP1 DIRECTS REPAIR TOWARD NHEJ AND AWAY FROM HR

In RNF168-deficient cells, 53BP1 foci are abolished and inefficient repair of DSBs during V(D)J recombination and CSR are observed, emphasizing the importance of RNF168 for 53BP1 recruitment and repair of RAG- and AID-induced DSBs. Since V(D)J recombination and CSR require repair of distal DNA breaks, the role of 53BP1 in promoting NHEJ has been investigated.

In 53BP1 knockout mice, long-range V(D)J recombination is impaired, but short-range recombination between D–J segments is not defective, supporting previous observations of increased frequency of short-range intra-switch recombination during CSR in the absence of 53BP1 (68). In addition to repair defects, 53BP1-deficient cells also exhibit end resection of unpaired V(D)J recombination-induced coding ends (69). Inhibiting DNA end resection is important for NHEJ-mediated repair of DSBs as the resulting ssDNA presents microhomologies for repair by alternative end joining (A-EJ) and HR pathways. Taken together, 53BP1 recruitment to DSB sites presents alternative, though not mutually exclusive, roles for repair during V(D)J recombination and CSR, as discussed below.

In a distance-dependent manner, 53BP1 mediates synopsis of distal DNA ends (**Figure 5**). It has been shown that 53BP1 is





capable of joining breaks between 1.2 and 96 kb long, the range of repair during rearrangement in V(D)J recombination and CSR (70). This distance-dependent repair function corresponds to  $\gamma$  H2AX spreading and is RNF8/RNF168-dependent. 53BP1 also has been shown to associate with motor proteins, possibly to enhance chromatin mobility during repair of distal DNA breaks (70, 71). In a distance-independent manner, 53BP1 is important for DNA end protection (71, 72). 53BP1 has been shown to block access of DNA nucleases to DNA ends, possibly due to its ability to constitutively interact with H4K20me2 (59, 70, 72). End protection is vital for V(D)J recombination and CSR as it prevents formation of ssDNA microhomologies that would permit repair by A-EJ or HR repair pathways. In this way, 53BP1 directs the cell to repair DSBs by NHEJ, shunning A-EJ or HR, and foci formation is dependent upon RNF168 ubiquitin ligase activity (Figure 5).

### CONCLUDING REMARKS

While programmed DNA breaks in Ig loci during B-cell development highlights the diverse repertoire of post-translation ubiquitin modifications, ubiquitination processes conversely have shed light on the vital aspects of protein regulation. Though normally avoided, DNA breaks are essential during B-cell development and for a functional adaptive immune system. Regulating the proteins and repair pathways involved during these processes is

therefore of utmost importance. RAG and AID off-target activity can be detrimental to the cell, introducing deleterious mutations and/or translocations with oncogenes. E3 ligases and ubiquitination events have been shown in the context of RAG regulation, AID activity and targeting, and recruitment of NHEJ repair proteins. Ubiquitination of RAG and AID are proposed to be important in restricting DNA damage activity at undesirable loci or during incorrect cell-cycle phase. This demonstrates roles for ubiquitin beyond canonical protein degradation. As ubiquitination is involved in protein stabilization and accumulation, recruitment, and co-factor binding, the cell is presented with another layer of regulation, the ubiquitin code.

### ACKNOWLEDGMENTS

This work was supported by grants from NIH Office of Director (1DP2OD008651-01), NIAID (1R01AI099195-01A1), and the Irma T. Hirsch Charitable Trust.

### REFERENCES

- Alt FW, Zhang Y, Meng FL, Guo C, Schwer B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* (2013) 152:417–29. doi:10.1016/j.cell.2013.01.007
- Bednarski JJ, Sleckman BP. Lymphocyte development: integration of DNA damage response signaling. *Adv Immunol* (2012) 116:175–204. doi:10.1016/B978-0-12-394300-2.00006-5

3. Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol* (2011) **11**:251–63. doi:10.1038/nri2941
4. Keim C, Kazadi D, Rothschild G, Basu U. Regulation of AID, the B-cell genome mutator. *Genes Dev* (2013) **27**:1–17. doi:10.1101/gad.200014.112
5. Xu Z, Zan H, Pone EJ, Mai T, Casali P. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat Rev Immunol* (2012) **12**:517–31. doi:10.1038/nri3216
6. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem* (2007) **76**:1–22. doi:10.1146/annurev.biochem.76.061705.090740
7. Komander D, Rape M. The ubiquitin code. *Annu Rev Biochem* (2012) **81**:203–29. doi:10.1146/annurev-biochem-060310-170328
8. Nishana M, Raghavan SC. Role of recombination activating genes in the generation of antigen receptor diversity and beyond. *Immunology* (2012) **137**:271–81. doi:10.1111/imm.12009
9. Swanson PC. The bounty of RAGs: recombination signal complexes and reaction outcomes. *Immunol Rev* (2004) **200**:90–114. doi:10.1111/j.0105-2896.2004.00159.x
10. McMahan CJ, Sadofsky MJ, Schatz DG. Definition of a large region of RAG1 that is important for coimmunoprecipitation of RAG2. *J Immunol* (1997) **158**:2202–10.
11. Jones JM, Gellert M. Autoubiquitylation of the V(D)J recombinase protein RAG1. *Proc Natl Acad Sci U S A* (2003) **100**:15446–51. doi:10.1073/pnas.2637012100
12. Kassmeier MD, Mondal K, Palmer VL, Raval P, Kumar S, Perry GA, et al. VprBP binds full-length RAG1 and is required for B-cell development and V(D)J recombination fidelity. *EMBO J* (2012) **31**:945–58. doi:10.1038/emboj.2011.455
13. Grazini U, Zanardi F, Citterio E, Casola S, Goding CR, McBlane F. The RING domain of RAG1 ubiquitylates histone H3: a novel activity in chromatin-mediated regulation of V(D)J joining. *Mol Cell* (2010) **37**:282–93. doi:10.1016/j.molcel.2009.12.035
14. Simkus C, Bhattacharyya A, Zhou M, Veenstra TD, Jones JM. Correlation between recombination activating gene 1 ubiquitin ligase activity and V(D)J recombination. *Immunology* (2009) **128**:206–17. doi:10.1111/j.1365-2567.2009.03101.x
15. Jones JM, Bhattacharyya A, Simkus C, Vallieres B, Veenstra TD, Zhou M. The RAG1 V(D)J recombinase/ubiquitin ligase promotes ubiquitylation of acetylated, phosphorylated histone 3.3. *Immunol Lett* (2011) **136**:156–62. doi:10.1016/j.imlet.2011.01.005
16. Yurchenko V, Xue Z, Sadofsky M. The RAG1 N-terminal domain is an E3 ubiquitin ligase. *Genes Dev* (2003) **17**:581–5. doi:10.1101/gad.1058103
17. Schatz DG, Swanson PC. V(D)J recombination: mechanisms of initiation. *Annu Rev Genet* (2011) **45**:167–202. doi:10.1146/annurev-genet-110410-132552
18. Villa A, Sobacchi C, Notarangelo LD, Bozzi F, Abinun M, Abrahamsen TG, et al. V(D)J recombination defects in lymphocytes due to RAG mutations: severe immunodeficiency with a spectrum of clinical presentations. *Blood* (2001) **97**:81–8. doi:10.1182/blood.V97.1.81
19. Simkus C, Anand P, Bhattacharyya A, Jones JM. Biochemical and folding defects in a RAG1 variant associated with Omenn syndrome. *J Immunol* (2007) **179**:8332–40.
20. Simkus C, Makiya M, Jones JM. Karyopherin alpha 1 is a putative substrate of the RAG1 ubiquitin ligase. *Mol Immunol* (2009) **46**:1319–25. doi:10.1016/j.molimm.2008.11.009
21. Lee J, Desiderio S. Cyclin A/CDK2 regulates V(D)J recombination by coordinating RAG-2 accumulation and DNA repair. *Immunity* (1999) **11**:771–81. doi:10.1016/S1074-7613(00)80151-X
22. Jiang H, Chang FC, Ross AE, Lee J, Nakayama K, Desiderio S. Ubiquitylation of RAG-2 by Skp2-SCF links destruction of the V(D)J recombinase to the cell cycle. *Mol Cell* (2005) **18**:699–709. doi:10.1016/j.molcel.2005.05.011
23. Mizuta R, Mizuta M, Araki S, Kitamura D. RAG2 is down-regulated by cytoplasmic sequestration and ubiquitin-dependent degradation. *J Biol Chem* (2002) **277**:41423–7. doi:10.1074/jbc.M206605200
24. Jiang H, Chang FC, Ross AE, Lee J, Nakayama K, Nakayama K, et al. Ubiquitylation of RAG-2 by Skp2-SCF links destruction of the V(D)J recombinase to the cell cycle. *Mol Cell* (2005) **18**:699–709. doi:10.1016/j.molcel.2005.05.011
25. Pillai S, Cariappa A. The follicular versus marginal zone B lymphocyte cell fate decision. *Nat Rev Immunol* (2009) **9**:767–77. doi:10.1038/nri2656
26. Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. *Nat Rev Immunol* (2008) **8**:22–33. doi:10.1038/nri2217
27. Peled JU, Kuang FL, Iglesias-Ussel MD, Roa S, Kalis SL, Goodman ME, et al. The biochemistry of somatic hypermutation. *Annu Rev Immunol* (2008) **26**:481–511. doi:10.1146/annurev.immunol.26.021607.090236
28. Odegard VH, Schatz DG. Targeting of somatic hypermutation. *Nat Rev Immunol* (2006) **6**:573–83. doi:10.1038/nri1896
29. Di Noia J, Neuberger MS. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* (2002) **419**:43–8. doi:10.1038/nature00981
30. Neuberger MS. Antibody diversification by somatic mutation: from Burnet onwards. *Immunol Cell Biol* (2008) **86**:124–32. doi:10.1038/sj.icb.7100160
31. Daniel JA, Nussenzweig A. The AID-induced DNA damage response in chromatin. *Mol Cell* (2013) **50**:309–21. doi:10.1016/j.molcel.2013.04.017
32. Yamane A, Resch W, Kuo N, Kuchen S, Li Z, Sun HW, et al. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat Immunol* (2011) **12**:62–9. doi:10.1038/ni.1964
33. Shen HM, Tanaka A, Bozek G, Nicolae D, Storb U. Somatic hypermutation and class switch recombination in Msh6(-/-)Ung(-/-) double-knockout mice. *J Immunol* (2006) **177**:5386–92.
34. Neuberger MS, Di Noia JM, Beale RC, Williams GT, Yang Z, Rada C. Somatic hypermutation at A.T pairs: polymerase error versus dUTP incorporation. *Nat Rev Immunol* (2005) **5**:171–8. doi:10.1038/nri1553
35. Yamane A, Robbiani DF, Resch W, Bothmer A, Nakahashi H, Oliveira T, et al. RPA accumulation during class switch recombination represents 5'-3' DNA-end resection during the S-G2/M phase of the cell cycle. *Cell Rep* (2013) **3**:138–47. doi:10.1016/j.celrep.2012.12.006
36. Xue K, Rada C, Neuberger MS. The in vivo pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in msh2-/- mice. *J Exp Med* (2006) **203**:2085–94. doi:10.1084/jem.20061067
37. Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill JC, et al. Proteasomal degradation restricts the nuclear lifespan of AID. *J Exp Med* (2008) **205**:1357–68. doi:10.1084/jem.20070950
38. Delker RK, Zhou Y, Strikoudis A, Stebbins CE, Papavasiliou FN. Solubility-based genetic screen identifies RING finger protein 126 as an E3 ligase for activation-induced cytidine deaminase. *Proc Natl Acad Sci U S A* (2013) **110**:1029–34. doi:10.1073/pnas.1214538110
39. Pavri R, Nussenzweig MC. AID targeting in antibody diversity. *Adv Immunol* (2011) **110**:1–26. doi:10.1016/B978-0-12-387663-8.00005-3
40. Kenter AL. AID targeting is dependent on RNA polymerase II pausing. *Semin Immunol* (2012) **24**:281–6. doi:10.1016/j.smim.2012.06.001
41. Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* (2012) **13**:720–31. doi:10.1038/nrg3293
42. Richard P, Manley JL. Transcription termination by nuclear RNA polymerases. *Genes Dev* (2009) **23**:1247–69. doi:10.1101/gad.1792809
43. Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, et al. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* (2010) **143**:122–33. doi:10.1016/j.cell.2010.09.017
44. Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, et al. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* (2011) **144**:353–63. doi:10.1016/j.cell.2011.01.001
45. Wang L, Whang N, Wuerffel R, Kenter AL. AID-dependent histone acetylation is detected in immunoglobulin S regions. *J Exp Med* (2006) **203**:215–26. doi:10.1084/jem.20051774
46. Peters A, Storb U. Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity* (1996) **4**:57–65. doi:10.1016/S1074-7613(00)80298-8
47. Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, et al. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell* (2011) **147**:95–106. doi:10.1016/j.cell.2011.07.048
48. Sun J, Rothschild G, Pefanis E, Basu U. Transcriptional stalling in B-lymphocytes: a mechanism for antibody diversification and maintenance of genomic integrity. *Transcription* (2013) **4**:127–35. doi:10.4161/trns.24556
49. Sun J, Keim CD, Wang J, Oliver P, Rabadan R, Basu U. E3-ubiquitin ligase Nedd4 determines the fate of AID-associated RNA polymerase II in B cells. *Genes Dev* (2013) **27**:1821–33. doi:10.1101/gad.210211.112
50. Goodarzi AA, Jeggo PA. The repair and signaling responses to DNA double-strand breaks. *Adv Genet* (2013) **82**:1–45. doi:10.1016/B978-0-12-407676-1.00001-9

51. Panier S, Durocher D. Regulatory ubiquitylation in response to DNA double-strand breaks. *DNA Repair (Amst)* (2009) **8**:436–43. doi:10.1016/j.dnarep.2009.01.013
52. Stucki M, Clapperton JA, Mohammad D, Yaffe MB, Smerdon SJ, Jackson SP. MDC1 directly binds phosphorylated histone H2AX to regulate cellular responses to DNA double-strand breaks. *Cell* (2005) **123**:1213–26. doi:10.1016/j.cell.2005.09.038
53. Kolas NK, Chapman JR, Nakada S, Ylanko J, Chahwan R, Sweeney FD, et al. Orchestration of the DNA-damage response by the RNF8 ubiquitin ligase. *Science* (2007) **318**:1637–40. doi:10.1126/science.1150034
54. Bekker-Jensen S, Lukas C, Kitagawa R, Melander F, Kastan MB, Bartek J, et al. Spatial organization of the mammalian genome surveillance machinery in response to DNA strand breaks. *J Cell Biol* (2006) **173**:195–206. doi:10.1083/jcb.200510130
55. Al-Hakim A, Escribano-Diaz C, Landry MC, O'Donnell L, Panier S, Szilard RK, et al. The ubiquitous role of ubiquitin in the DNA damage response. *DNA Repair (Amst)* (2010) **9**:1229–40. doi:10.1016/j.dnarep.2010.09.011
56. Durocher D, Henckel J, Fersht AR, Jackson SP. The FHA domain is a modular phosphopeptide recognition motif. *Mol Cell* (1999) **4**:387–94. doi:10.1016/S1097-2765(00)80340-8
57. Mailand N, Bekker-Jensen S, Fastrup H, Melander F, Bartek J, Lukas C, et al. RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell* (2007) **131**:887–900. doi:10.1016/j.cell.2007.09.040
58. Huen MS, Grant R, Manke I, Minn K, Yu X, Yaffe MB, et al. RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. *Cell* (2007) **131**:901–14. doi:10.1016/j.cell.2007.09.041
59. Botuyan MV, Lee J, Ward IM, Kim JE, Thompson JR, Chen J, et al. Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell* (2006) **127**:1361–73. doi:10.1016/j.cell.2006.10.043
60. Doil C, Mailand N, Bekker-Jensen S, Menard P, Larsen DH, Pepperkok R, et al. RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins. *Cell* (2009) **136**:435–46. doi:10.1016/j.cell.2008.12.041
61. Stewart GS, Panier S, Townsend K, Al-Hakim AK, Kolas NK, Miller ES, et al. The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage. *Cell* (2009) **136**:420–34. doi:10.1016/j.cell.2008.12.042
62. Mattioli F, Vissers JH, van Dijk WJ, Ikpa P, Citterio E, Vermeulen W, et al. RNF168 ubiquitinates K13–15 on H2A/H2AX to drive DNA damage signaling. *Cell* (2012) **150**:1182–95. doi:10.1016/j.cell.2012.08.005
63. Panier S, Ichijima Y, Fradet-Turcotte A, Leung CC, Kaustov L, Arrowsmith CH, et al. Tandem protein interaction modules organize the ubiquitin-dependent response to DNA double-strand breaks. *Mol Cell* (2012) **47**:383–95. doi:10.1016/j.molcel.2012.05.045
64. Stolz A, Ertych N, Bastians H. Tumor suppressor CHK2: regulator of DNA damage response and mediator of chromosomal stability. *Clin Cancer Res* (2011) **17**:401–5. doi:10.1158/1078-0432.CCR-10-1215
65. Yamanaka K, Dorjsuren D, Eoff RL, Egli M, Maloney DJ, Jadhav A, et al. A comprehensive strategy to discover inhibitors of the translesion synthesis DNA polymerase kappa. *PLoS One* (2012) **7**:e45032. doi:10.1371/journal.pone.0045032
66. Acs K, Luijsterburg MS, Ackermann L, Salomons FA, Hoppe T, Dantuma NP. The AAA-ATPase VCP/p97 promotes 53BP1 recruitment by removing L3MBTL1 from DNA double-strand breaks. *Nat Struct Mol Biol* (2011) **18**:1345–50. doi:10.1038/nsmb.2188
67. Meerang M, Ritz D, Paliwal S, Garajova Z, Bosshard M, Mailand N, et al. The ubiquitin-selective segregase VCP/p97 orchestrates the response to DNA double-strand breaks. *Nat Cell Biol* (2011) **13**:1376–82. doi:10.1038/ncb2367
68. Reina-San-Martin B, Chen J, Nussenzweig A, Nussenzweig MC. Enhanced intra-switch region recombination during immunoglobulin class switch recombination in 53BP1-/- B cells. *Eur J Immunol* (2007) **37**:235–9. doi:10.1002/eji.200636789
69. Difilippantonio S, Gapud E, Wong N, Huang CY, Mahowald G, Chen HT, et al. 53BP1 facilitates long-range DNA end-joining during V(D)J recombination. *Nature* (2008) **456**:529–33. doi:10.1038/nature07476
70. Bothmer A, Robbiani DF, Di Virgilio M, Bunting SF, Klein IA, Feldhahn N, et al. Regulation of DNA end joining, resection, and immunoglobulin class switch recombination by 53BP1. *Mol Cell* (2011) **42**:319–29. doi:10.1016/j.molcel.2011.03.019
71. Dimitrova N, Chen YC, Spector DL, de Lange T. 53BP1 promotes non-homologous end joining of telomeres by increasing chromatin mobility. *Nature* (2008) **456**:524–8. doi:10.1038/nature07433
72. Bunting SF, Callen E, Wong N, Chen HT, Polato F, Gunn A, et al. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* (2010) **141**:243–54. doi:10.1016/j.cell.2010.03.012
73. Bohgaki T, Bohgaki M, Cardoso R, Panier S, Zeegers D, Li L, et al. Genomic instability, defective spermatogenesis, immunodeficiency, and cancer in a mouse model of the RIDDLE syndrome. *PLoS Genet* (2011) **7**:e1001381. doi:10.1371/journal.pgen.1001381

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 December 2013; accepted: 25 February 2014; published online: 11 March 2014.

Citation: Chao J, Rothschild G and Basu U (2014) Ubiquitination events that regulate recombination of immunoglobulin loci gene segments. *Front. Immunol.* **5**:100. doi: 10.3389/fimmu.2014.00100

This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Chao, Rothschild and Basu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Function of YY1 in long-distance DNA interactions

Michael L. Atchison\*

Department of Animal Biology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA, USA

**Edited by:**

Ananda L. Roy, Tufts University School of Medicine, USA

**Reviewed by:**

Yoshiteru Sasaki, Kyoto University, Japan

Masaki Hikida, Kyoto University, Japan

**\*Correspondence:**

Michael L. Atchison, Department of Animal Biology, School of Veterinary Medicine, University of Pennsylvania, 3800 Spruce Street, Philadelphia, PA 19104, USA  
e-mail: atchison@vet.upenn.edu

During B cell development, long-distance DNA interactions are needed for V(D)J somatic rearrangement of the immunoglobulin (Ig) loci to produce functional Ig genes, and for class switch recombination (CSR) needed for antibody maturation. The tissue-specificity and developmental timing of these mechanisms is a subject of active investigation. A small number of factors are implicated in controlling Ig locus long-distance interactions including Pax5, Yin Yang 1 (YY1), EZH2, IKAROS, CTCF, cohesin, and condensin proteins. Here we will focus on the role of YY1 in controlling these mechanisms. YY1 is a multifunctional transcription factor involved in transcriptional activation and repression, X chromosome inactivation, Polycomb Group (PcG) protein DNA recruitment, and recruitment of proteins required for epigenetic modifications (acetylation, deacetylation, methylation, ubiquitination, sumoylation, etc.). YY1 conditional knock-out indicated that YY1 is required for B cell development, at least in part, by controlling long-distance DNA interactions at the immunoglobulin heavy chain and Igk loci. Our recent data show that YY1 is also required for CSR. The mechanisms implicated in YY1 control of long-distance DNA interactions include controlling non-coding antisense RNA transcripts, recruitment of PcG proteins to DNA, and interaction with complexes involved in long-distance DNA interactions including the cohesin and condensin complexes. Though common rearrangement mechanisms operate at all Ig loci, their distinct temporal activation along with the ubiquitous nature of YY1 poses challenges for determining the specific mechanisms of YY1 function in these processes, and their regulation at the tissue-specific and B cell stage-specific level. The large numbers of post-translational modifications that control YY1 functions are possible candidates for regulation.

**Keywords: YY1, polycomb, condensin, cohesin, DNA loops, immunoglobulin loci**

## THE EARLY DAYS

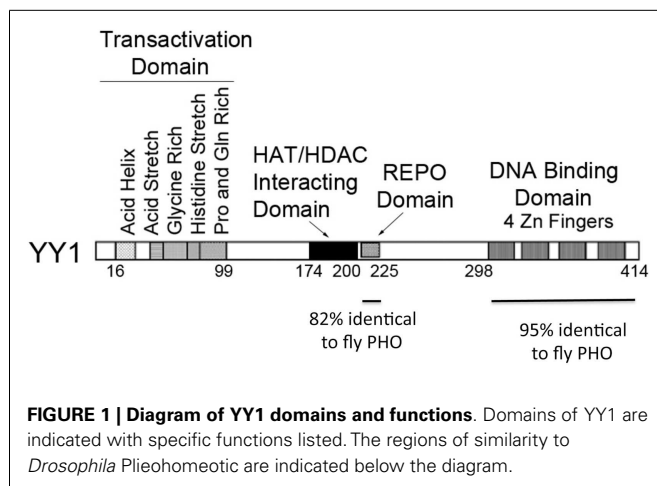
Yin Yang 1 (YY1) was first identified in 1985 as a factor that yielded an *in vivo* B cell-specific DMS methylation interference pattern over the immunoglobulin heavy chain (IgH) intron enhancer (1, 2). The enhancer site that bound YY1 was defined as the  $\mu$ E1 site (3) and nuclear factors that bound to this sequence were identified by EMSA (4). Our laboratory isolated a cDNA clone expressing a protein that bound to the Igk3' enhancer as well as the IgH  $\mu$ E1 site and named the protein NF-E1 (5). Simultaneously the factor was cloned by Tom Shenk's laboratory and named YY1 (6) based on its ability to bind the adenoviral P1 promoter and both activate and repress transcription, by Robert Perry's laboratory and named delta (7) due to its binding to the delta motif in the promoters of ribosomal protein genes, and by Keiko Ozato's laboratory and named UCRBP based on its ability to bind to the upstream control region of retroviral LTRs (8). Ultimately, the name YY1 was adopted by all.

Yin Yang 1 contains four zinc fingers at its carboxyl terminus (amino acids 298–414) and a region rich in alanine and glycine between amino acids 154 and 201. The first 100 amino acids of YY1 encode several notable features. Sequences 43–53 contain 11 consecutive acidic residues while amino acids 70–80 consist of 11 consecutive histidine residues. These two segments are separated by a region rich in glycine (residues 54–69). In addition, sequences 16–29 have the potential to form an amphipathic

negatively charged helix and sequences 80–100 are rich in proline and glutamine. Sequences near the carboxyl terminus (333–397), which overlap the YY1 zinc fingers, and sequences 170–200 have been reported to be involved in transcriptional repression (6, 9–15). These sequences are known to physically interact with a variety of transcriptionally important proteins including TBP, p300, c-myc, and HDAC2 (16). YY1 sequences important for transcriptional activation reside near the amino-terminus (9, 12, 13, 17). **Figure 1** shows various sequence features and functional domains of YY1.

## DIVERSE AND COMPLEX ROLES OF YY1

Over the past 22 years, multiple diverse YY1 functions have been identified. YY1 is crucial for embryonic development because homozygous mutation of the *yy1* gene in mice results in peri-implantation lethality (18). YY1 is implicated in lineage differentiation of skeletal and cardiac muscle, and in cell growth control (13, 17, 19–24), as well as disease pathways such as dystrophic muscle disease (25–27). YY1 and its target genes are also believed to be central regulators of germinal center B cell development (28), and YY1 has been suggested to regulate genomic targeting of activation induced cytidine deaminase (AID) (29). YY1 is implicated in a number of cancers (30–32), and is overexpressed in B cell lymphomas that depend on AID function. YY1 is associated with B cell transformation and tumor progression in diffuse



large B cell lymphoma (DLBCL) (33, 34), and high levels of YY1 expression are associated with reduced patient survival in DLBCL as well as follicular lymphoma. CTCF-YY1 elements are clustered in the imprinting domain of Tsix (35) and YY1 docks Xist particles on the X chromosome via DNA and RNA interactions during X chromosome inactivation (36). YY1 can also control imprinting at the Peg3 and Gnas domains (37). YY1 can control human immunodeficiency virus (HIV) gene expression and viral titers, and deletion of YY1 binding sites in regulatory regions of human papilloma viruses correlates with increased viral gene expression and the development of cervical cancer (38–46). Thus, YY1 function is related to transcriptional regulation, embryonic development, X-chromosome inactivation, imprinting, oncogenesis, viral gene expression, epigenetic function, and a growing list of diseases.

### IDENTIFICATION OF THE PcG FUNCTION OF YY1

A significant new function of YY1 was suggested in 1998 when the Kassis laboratory cloned the *Drosophila* Pleiohomeotic (PHO) sequence and observed similarity to YY1 (47) (Figure 1). Girtton and Jeon (48) demonstrated that PHO is a Polycomb Group (PcG) protein, a family of proteins involved in epigenetic chromosomal condensation, stable transcriptional repression, control of cell proliferation, hematopoietic development, as well as stem cell self-renewal. This raised the exciting possibility that YY1 is a vertebrate PcG protein. PHO is highly homologous to YY1 in two regions. These two regions include YY1 sequences 296–414 and 205–226 (the corresponding segments in PHO are residues 357–475 and 148–169, respectively). Sequences 298–414 constitute the four YY1 zinc fingers. The homology over this region is extraordinary for organisms as diverse as flies and humans (112 identities out of 118; 95%). Within this segment, zinc fingers 2 and 3 are 100% identical. The 205–226 segment is also highly homologous (18/22; 82% identity). Outside of these regions of high similarity, YY1 and PHO showed no discernible similarity. PHO does not contain an obvious transcriptional activation domain and lacks YY1 structural features such as acid and histidine stretches. However, the two regions of high similarity between YY1 and PHO, and

their similar spatial locations within the proteins, suggested that they might carry out some of the same functions in vertebrates and flies, respectively.

Prompted by the possibility that YY1 functions as a PcG protein, we tested this hypothesis using a *Drosophila* *in vivo* transcription system, as well as a phenotypic correction assay. Our results showed that human YY1 does indeed function as a PcG protein *in vivo* (49–51). We found that YY1 can repress transcription in a PcG-dependent fashion, can phenotypically correct *pho* mutant flies, and can recruit PcG proteins to specific DNA sequences resulting in tri-methylation of H3 lysine 27 (49–51). The mechanisms responsible for targeting mammalian PcG proteins to specific DNA regions has long been proven enigmatic because none of the components of the PcG complexes bind to specific DNA sequences, yet the PcG complexes associate with specific DNA regions *in vivo*. Our demonstration that YY1 is a mammalian PcG protein with high affinity sequence-specific DNA binding activity suggested that YY1 is a crucial factor for targeting specific proteins to specific DNA sequences. The role of YY1 in PcG targeting has been confirmed in a number of studies (52–55) though clearly other factors are involved as YY1 (and PHO) does not co-localize with PcG proteins in all cell types (56–58). A particularly exciting aspect of YY1 PcG function is that PcG proteins are known to contribute to B cell development, and the PcG protein EZH2, like YY1, is required for Ig locus contraction (further explained below) (59). Nucleation of PcG proteins to specific target DNA sites by YY1 within the Ig loci thus opens up a new avenue for mechanistic evaluation of B cell development and Ig locus contraction, because PcG proteins are capable of mediating long-distance DNA interactions (60).

### THE YY1 REPO DOMAIN

Using a fly transgenic approach, we set out to identify the YY1 sequences involved in PcG function (61). We found that the region of 82% YY1-PHO identity (the 25 amino acids between residues 201 and 226), when fused to a heterologous GAL4 DNA binding domain, was necessary and sufficient for PcG-dependent transcriptional repression. Amazingly, this small 25 amino acid segment was also necessary and sufficient for recruitment of PcG proteins to DNA resulting in tri-methylation of H3 lysine 27. Therefore, we named YY1 sequences 201–226 the REPO domain for their ability to REcruit Polycomb (61). A REPO domain YY1 mutant ( $\Delta$ 201–226) can mediate nearly all YY1 functions such as DNA binding, transcriptional activation, transient transcriptional repression, and interaction with HDAC proteins. However, this mutant fails to carry out YY1 PcG functions and fails to recruit PcG proteins to DNA (61). How the YY1 REPO domain recruits PcG proteins to DNA is now being elucidated. Two homologous proteins, YAF2 and RYBP, were previously identified as YY1 interacting proteins (62, 63). Functionally, RYBP associates with a subset of PcG complexes named PRC1L4 (64) and is involved in the repressive function of *hoxD11.12*, a mammalian “PRE-like” sequence (65). YAF2 was first identified by its ability to bind to YY1 (63) and we found YAF2 can interact with the REPO domain perhaps functioning as a bridge protein in PcG recruitment (52, 66). The importance of the YY1 REPO domain for B cell development is discussed below.

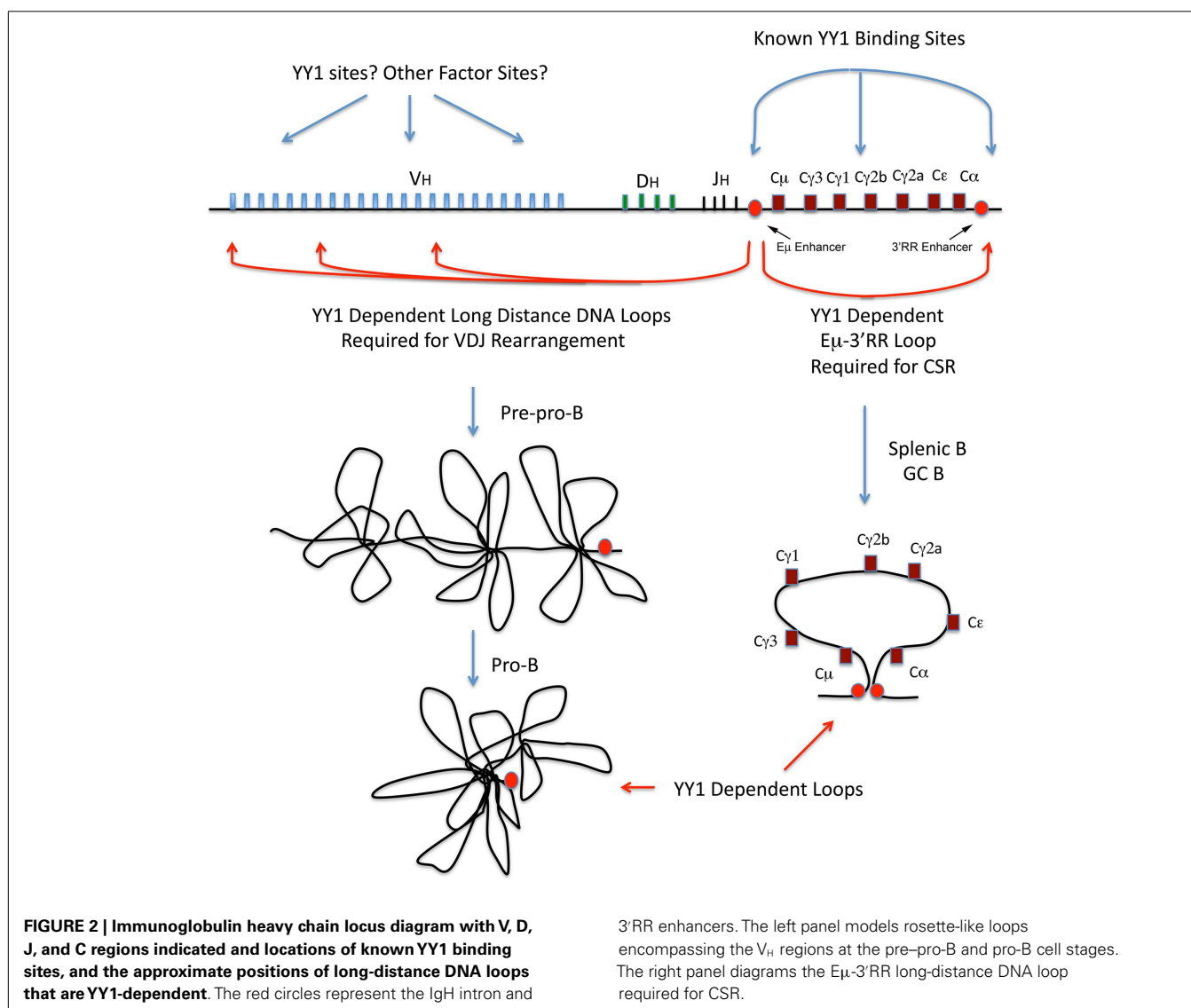
## STRUCTURE OF IMMUNOGLOBULIN LOCI DURING B CELL DEVELOPMENT

B cell development involves progression from Lin<sup>-</sup> Sca-1<sup>+</sup> c-kit<sup>+</sup> (LSK) progenitor cells through a number of intermediate B cell stages including pro-B, pre-B, immature B, mature B, and plasma cell stages. The early stages of B cell development can be delineated by the rearrangement status of the immunoglobulin heavy and light chain genes. Both heavy and light chain genes are produced during early, antigen-independent B cell development by a somatic rearrangement process that links together either V, D, and J segments (heavy chain), or V and J segments (light chain) to produce functional Ig genes (67–70). The Ig loci are huge (2.4–3.2 Mb) and for rearrangement of distal variable region genes to occur, the loci must go through a physical contraction process. Prior to the onset of rearrangement, Ig loci reside at the nuclear periphery in an “extended” configuration. However, at the pro-B cell stage, when the heavy chain genes undergo rearrangement, the loci take up an intranuclear localization with concomitant contraction of the loci (heavy chain first followed by light chain) (71–74). While IgH DJ

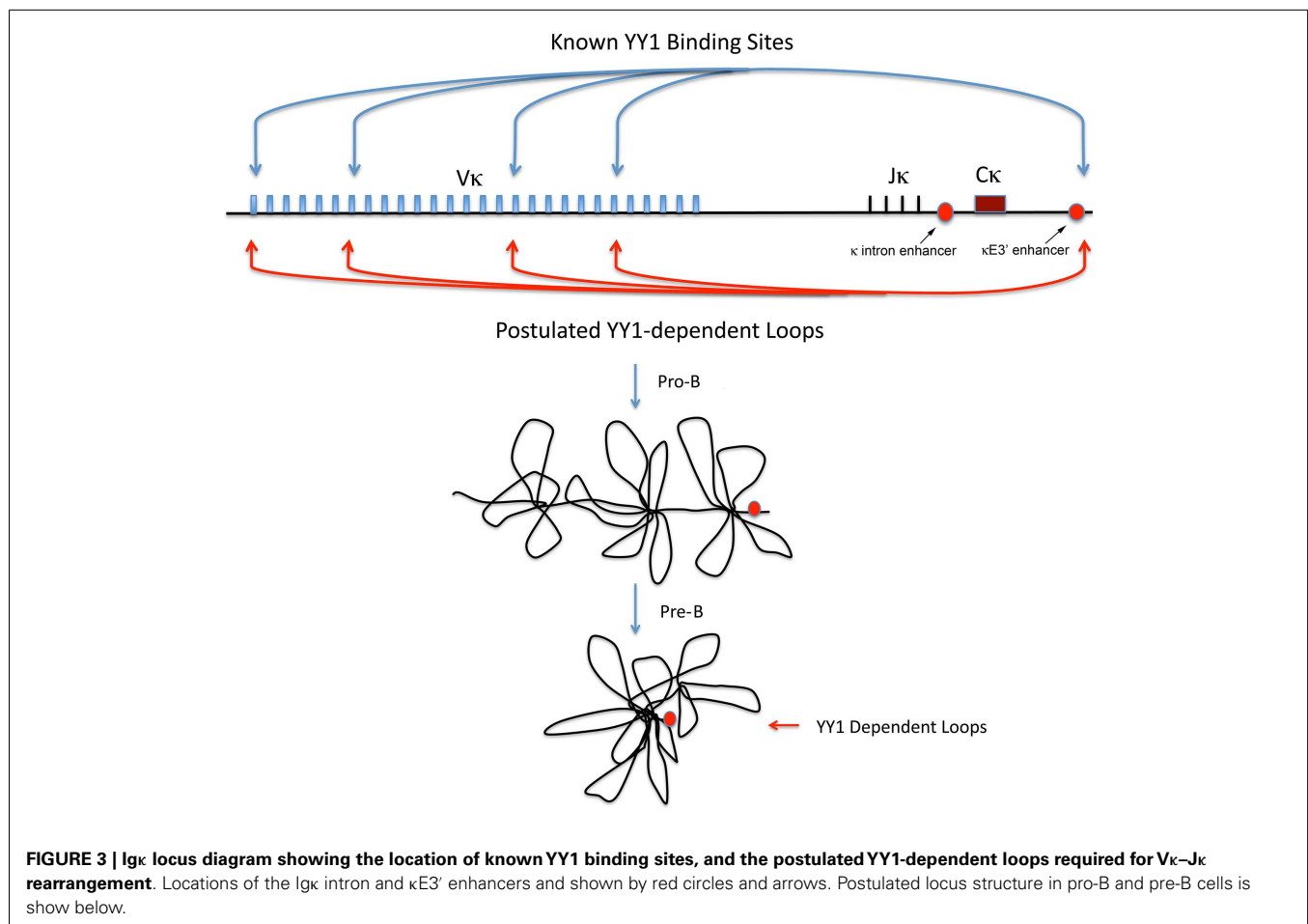
and proximal V<sub>H</sub> to D and V<sub>K</sub> to J<sub>K</sub> rearrangements can occur without contraction, the distal V genes require locus contraction and looping for rearrangement (71–73, 75–77).

Current data suggest that the Ig loci are organized as loops into rosette-like structures separated by spacer DNA (76, 78–80). A number of domains have been identified at the IgH locus, which adopt various conformations during development (76, 78–80). At the pre-pro-B cell stage, these rosette domains are in an extended conformation, but in pro-B cells the structure changes such that each V region domain is repositioned with all V<sub>H</sub> regions approximately equidistant to the D<sub>H</sub> and J<sub>H</sub> regions, thus affording roughly equal access for recombination (79, 80) (Figure 2, left panel). Similar structures are believed to exist at the Ig kappa locus at pro-B and pre-B cell stages (Figure 3).

The mechanisms that control Ig locus contraction are unknown. A small number of transcription factors or protein complexes (YY1, Pax5, CTCF, IKAROS, cohesin, condensin, EZH2) are implicated in the DNA loops needed for V(D)J rearrangement (59, 78, 81–86), but the molecular details and regulatory processes that



3'RR enhancers. The left panel models rosette-like loops encompassing the V<sub>H</sub> regions at the pre-pro-B and pro-B cell stages. The right panel diagrams the E<sub>μ</sub>-3'RR long-distance DNA loop required for CSR.



control this mechanism are not clear. Pax5 binds to multiple repeat sequences in the distal region of the IgH locus (PAIR sequences) and is believed to participate in rearrangement of distal V<sub>H</sub> genes (83). Non-coding antisense transcripts expressed across the PAIR sequences correlate with VDJ rearrangement and are postulated to be involved with IgH locus contraction (83, 87, 88). Pax5 controls some of these transcripts (83), and recently YY1 was shown to regulate antisense transcripts across at least two PAIR sequences (87). Many Pax5 and YY1 potential binding sites exist in the IgH locus (89) and these transcription factors co-localize at some of these sites (87). Similar to the Pax5 and YY1 knock-out phenotypes (discussed below), PcG protein EZH2 knock-out results in arrest at the pro-B cell stage with impaired distal V<sub>H</sub> to D<sub>H</sub>-J<sub>H</sub> rearrangement (59). CTCF and cohesin have been argued to regulate Ig locus structure and to control interactions of D<sub>H</sub> and J<sub>H</sub> regions with proximal V<sub>H</sub> segments and J<sub>κ</sub> regions with proximal V<sub>κ</sub> segments (81, 82, 90–92). Ikaros knock-out also impacts IgH rearrangement as well as locus contraction (93).

### THE ROLE OF YY1 AND THE REPO DOMAIN IN B CELL DEVELOPMENT

Yin Yang 1 has long been believed to play some role in immunoglobulin (Ig) gene regulation and B cell biology because it associates with multiple Ig enhancer elements including the

heavy chain intron and 3' enhancers, the Ig kappa 3' enhancer, as well as to a site between the C<sub>H</sub> γ1 and γ2b exons (1–5, 87, 94) (Figures 2 and 3). The Shi laboratory at Harvard provided insight into the role of YY1 in B cell development by demonstrating that conditional knock-out of YY1 in the B cell lineage (using mb1-CRE which is expressed early after B lineage commitment) resulted in arrest at the pro-B cell stage (84). Pro-B cells lacking YY1 have normal D<sub>H</sub>-J<sub>H</sub> recombination but reduced frequency of V<sub>H</sub>-D<sub>H</sub>-J<sub>H</sub> recombination, with the defect being most severe for more distal V<sub>H</sub> genes (84). These knock-out pro-B cells showed a defect in Ig locus contraction (84), and this phenotype has been confirmed by a number of studies (81, 88). Thus, conditional knock-out of YY1 using mb1-CRE results in arrest at the pro-B cell stage, lost Ig locus contraction, and reduced rearrangement of distal V genes. Importantly, despite the fact that proximal VDJ recombination does occur, very few mature B cells are generated in conditional knock-outs. Furthermore, introduction of a rearranged heavy chain gene only partially complements the YY1 conditional knock-out phenotype, suggesting additional roles for YY1 in early B cell development (84).

Intrigued by the similarity between the YY1 and PcG protein EZH2 B cell knock-out phenotypes (59, 84), we set out to determine the importance of YY1 PcG function for B cell development. Using YY1 wild-type and YY1ΔREPO retroviral constructs, we



transduced bone marrow from *yy1<sup>fl/fl</sup> mb1-CRE* mice and injected this transduced bone marrow into irradiated secondary recipients. Thus, within the B cell lineage of the transplanted mice, only the transduced YY1 constructs will provide YY1 function due to deletion of the endogenous *yy1* gene by *mb1-CRE* action. While wild-type YY1 largely restored B cell development, the YY1ΔREPO reconstituted cells arrested B cell development at the pro-B and pre-B cell stages (85). Interestingly, IgH VDJ rearrangement was largely normal, but Igk rearrangement showed a dramatically skewed repertoire. Only a small number of V<sub>k</sub> genes underwent rearrangement with one third of rearrangements to the most distal 5' V kappa gene. This dramatic result suggested that in the absence of YY1 PcG function, most of the DNA loops at the Igk locus needed for Igk rearrangement were abrogated, and a small number of loops that are independent of YY1 PcG function remained for Igk V<sub>k</sub>-J<sub>k</sub> rearrangements. At least some of these loops may require E2A or Pax5 (85), although this is speculative.

### MECHANISMS OF Ig LOCUS CONTRACTION

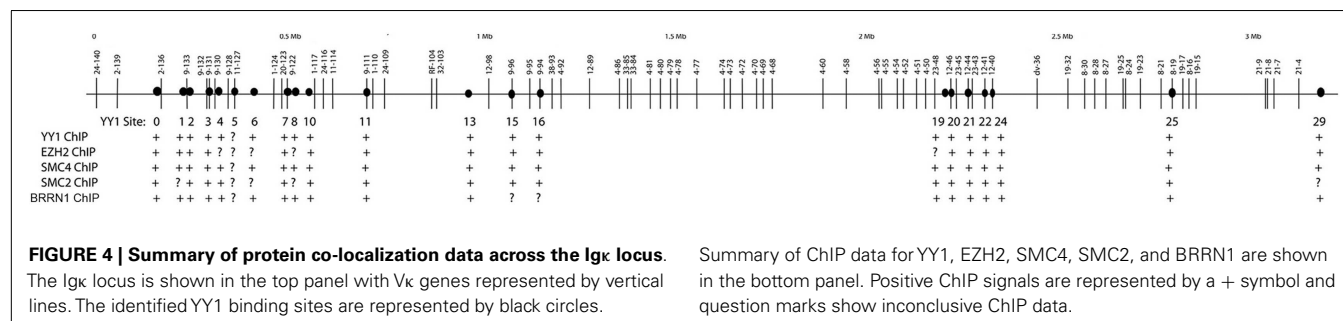
The dramatically skewed V<sub>k</sub>-J<sub>k</sub> rearrangement profiles in YY1ΔREPO compared to wild-type YY1 mice (85), suggested a possible direct effect of YY1 on Igk locus structure, and loss of IgH locus contraction in a YY1 knock-out background suggested parallel effects at the heavy chain locus. Consistent with a direct effect on Igk locus structure, RNAi knock-down of YY1 in bone marrow cultures reduced Igk rearrangement at a subset of V<sub>k</sub> genes (85). Since the Shi lab showed YY1 is important for Ig locus contraction (81, 84, 88), we hypothesized that clusters of YY1 binding sites exist across the Ig loci, and that YY1 binding to these sites would result in recruitment of proteins needed for Ig locus contraction. As predicted, we identified clusters of YY1 binding sites across the Igk locus that binds to YY1 (85). We found that PcG protein EZH2 co-localized with YY1 at these sites apparently as a result of recruitment by YY1 (85). We also identified several proteins that physically interact with the YY1 REPO domain providing potential insight into the mechanism of YY1 function in locus contraction. Intriguingly, we found that proteins from the condensin and cohesin complexes (SMC4 and SMC1) that are needed for contraction of chromosomes during mitosis (95–99), as well as lamin proteins, bind to the YY1 REPO domain. Lamin proteins are known to be involved in long-distance DNA interactions (100–103). Similarly, cohesin and condensin complexes, along with topoisomerase 2, are involved in mitotic chromosome contraction and higher order chromosome organization and dynamics (96, 104). During mitosis, condensin and cohesin proteins associate

with the chromosomes and function in chromosomal contraction, cohesion, assembly, and segregation (96–98). A subpopulation of these proteins remains chromosome-associated at specific foci in the interphase nucleus (98). Importantly, cohesin and condensin proteins are involved in numerous long-distance DNA interactions (92, 105–114). Therefore, we hypothesized condensin and cohesin proteins associate with Ig loci in pro-B and pre-B cells by virtue of interaction with YY1, and thereby function to participate in Ig locus contraction. Consistent with this idea, we found that condensin proteins associate with the clusters of YY1 binding sites that we identified within the Igk locus (85) in primary pro-B cells, but not in fibroblasts suggesting a B cell specific function of condensin and cohesin association with these sites (see **Figure 4**).

To test the functional consequences of YY1 and condensin binding at the Ig kappa locus, we performed RNAi knock-down and Ig kappa rearrangement assays. We found that knock-down of YY1 or condensin proteins resulted in reduced Igk rearrangement at a subset of V<sub>k</sub> genes (85). Thus, YY1 binds to sites in the Ig loci, perhaps recruits PcG, condensin, cohesin, and lamin proteins to these sites, and results in specific Ig locus chromosomal contraction. The identification of condensin mutants that specifically affect T cell development supports the idea of condensin proteins (which are ubiquitously expressed) having lymphoid specific functions (115). These complexes can mediate long-distance chromosomal interactions (105, 107), and kleisin-β, a member of the condensin II complex is important for T cell development as is cohesin subunit Rad21 (92, 115). Cohesin subunit Rad 21 (a kleisin family protein) is recruited to CTCF binding sites throughout the Ig loci during B lymphocyte development (82). As condensin I is involved in the process of physically compacting DNA in the presence of hydrolyzable ATP (116), condensin complex proteins may also participate in bringing V genes in the Ig locus into close proximity with D and J gene segments.

### LONG-DISTANCE DNA INTERACTIONS AND CSR

Long-distance DNA loops are also required for class switch recombination (CSR), which recombines the rearranged VDJ segments that provide antibody specificity with various Ig heavy chain constant (C) regions with different effector functions (117, 118). CSR requires a large 220 kb long-distance DNA loop synapse between the IgH intron enhancer (E<sub>μ</sub>) region, and the 3'RR enhancer downstream of the 3'-most C<sub>α</sub> exon (119, 120) (the E<sub>μ</sub>-3'RR synapse; see **Figure 2**, right panel). In addition, CSR to individual IgH C exons requires formation of inducible DNA loops from each switch region DNA sequence into the E<sub>μ</sub>-3'RR synapse (119,





120). Over 40 proteins are involved in the enzymology and mechanism of CSR and include DNA repair (base excision repair and mismatch repair) proteins, DNA damage sensors, factors that alter chromatin structure, factors that bind to AID, and transcriptional regulatory proteins [reviewed in Ref. (121)]. However, none of these factors are known to specifically impact the  $E\mu$ -3'RR DNA loop required for CSR.

Recent progress, however, has shed light on these long-distance DNA loops. CTCF and cohesin bind to the IgH 3'RR enhancer within the hs5–7 sites (81, 122, 123), and cohesin binding is induced at certain  $C_H$  switch regions in response to inducers of CSR implying a function for cohesin in CSR (123). Consistent with this, knock-down of cohesin subunits impairs CSR (123). In addition, knock-down of the cohesin loading protein NIPBL reduces CSR, reduces non-homologous end joining, and increases microhomology end joining (124). Interestingly, AID was shown to physically interact with condensin, cohesin, and INO80 complex proteins (123), precisely the same complexes that bind to YY1 (85, 125, 126).

Notably, we found that YY1 conditional knock-out in splenic B cells significantly reduces CSR (127). YY1 physically interacts with AID, leading to stabilization and increased AID nuclear accumulation, and this control of AID nuclear levels can regulate CSR. Control of nuclear levels of AID is crucial not only for regulating antibody maturation processes (CSR and somatic hypermutation), but also is important for maintaining integrity of the mammalian genome. Elevated levels of YY1 could cause aberrant accumulation of AID in germinal center B cells leading to increased mutagenesis and lymphomagenesis. Indeed, YY1 levels are elevated in germinal center-derived human DLBCL (34), suggesting that YY1 contributes to disease progression. However, we also found that YY1 has a second function important for CSR. In collaboration with Ranjan Sen (NIA), we found that YY1 is necessary for long-distance DNA loops formed between the  $E\mu$  and 3'RR enhancers (unpublished data). Recently, Kenter and colleagues identified a long-distance DNA loop between the  $E\mu$  and hs3b–hs4 sites of the 3'RR that is dramatically induced upon induction of CSR in splenic B cells (119). We found that this long-distance DNA loop is YY1-dependent (unpublished data). Thus, YY1 controls long-distance DNA loops in splenic B cells that are critical for CSR. Can the same be said of the long-distance DNA loops needed for IgH V(D)J rearrangement, and perhaps for other long-distance DNA loops? Recent evidence suggests this is the case.

### YY1-DEPENDENT IgH LONG-DISTANCE DNA INTERACTIONS

The Sen Laboratory and colleagues identified long-distance DNA loops in both the  $V_H$  distal and proximal regions, and at the 3' end of the locus (78). They found YY1 bound to many of these segments and postulated either homotypic YY1 interactions to mediate these loops, or heterotypic interactions with other proteins (78). The essential nature of YY1 for these loops was subsequently demonstrated. In pro-B cells, YY1 conditional knock-out ablates long-distance DNA loops between the  $E\mu$  region and the distal and proximal  $V_H$  regions (87). In addition, YY1 knock-out in pro-B cells ablates loops between the  $E\mu$  region and the 3'RR enhancer, hs5–7 region (87). Thus, YY1 is essential for long-distance DNA loops within the IgH locus involved in either VDJ rearrangement,

or CSR (Figure 2). Finally, YY1 is also involved in long-distance DNA interactions at the Th2 cytokine locus and controls IL4, IL5, and IL13 expression (128). These dramatic results indicate that YY1 is required for long-distance DNA loops that control IgH V(D)J rearrangement, CSR, and gene regulation. Our studies at the Igk locus (85) also indicate a role for YY1 in long-distance DNA interactions needed for Igk rearrangement (Figure 3).

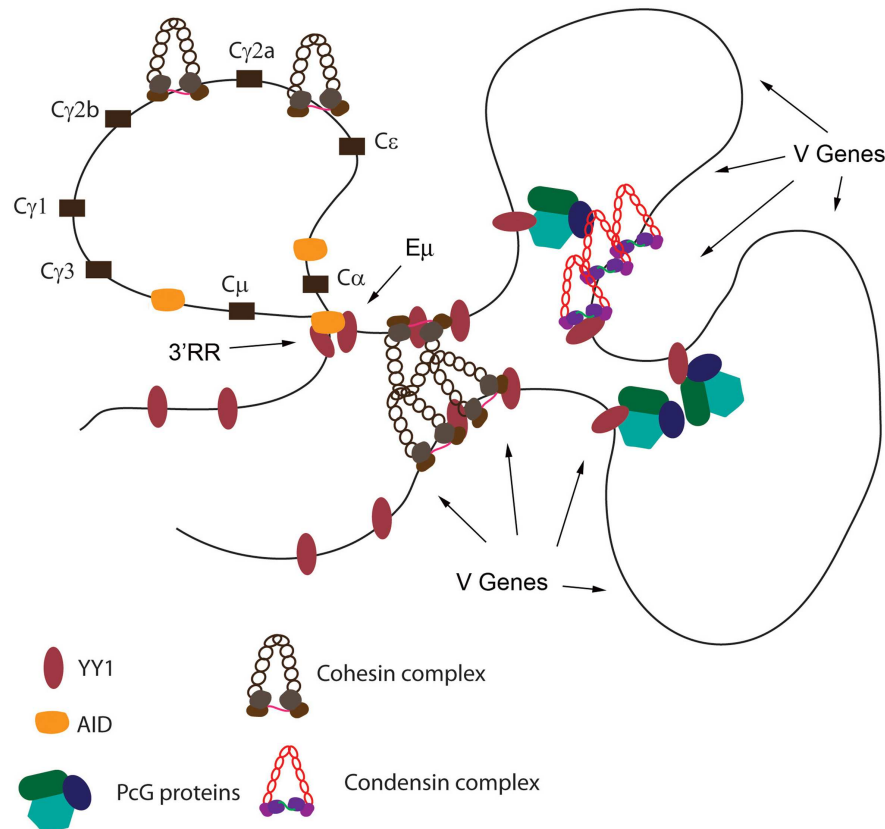
### REGULATORY MECHANISMS FOR YY1 FUNCTION

How might YY1 be functioning in these diverse long-distance DNA interactions? As described above, in pro-B cells, YY1 binds constitutively to the  $E\mu$  enhancer, to hs5–7 sites in the 3'RR enhancer, to a site between the Cy1 and Cy2b exons, and inducibly to the hs3b site in the 3'RR enhancer in splenic B cells (5, 78, 87, 94). The mechanism of regulation of developmental stage-specific function of YY1 in VDJ rearrangement at the IgH locus (pro-B cells), in  $V\kappa$ –Jk rearrangement at the Igk locus (pre-B cells), and in CSR at the IgH locus (mature splenic B cells), is presently unknown. YY1 may participate in regulatory stage-specific functions to control locus accessibility (129), but other factors may control accessibility enabling subsequent YY1 DNA binding.

Yin Yang 1 function can be regulated by a number of mechanisms. Stage-specific regulation could be at the level of YY1 DNA binding, such as the LPS inducible binding in the 3'RR enhancer in splenic B cells. YY1 binding to the Ig heavy chain 3'RR hypersensitive site 3b (hs3b) as well as to the  $E\mu$  enhancer is inducible by LPS (94). In this case, YY1 appears to be sequestered from DNA in resting B lymphocytes through interaction with hypophosphorylated retinoblastoma protein (Rb). However, after LPS induction, Rb becomes hyperphosphorylated and releases YY1 enabling it to bind to the hs3b and  $E\mu$  enhancers. Interestingly, hs3b and 4 hypersensitive sites are crucial for formation of  $E\mu$ :3'RR1 enhancer synapses with germline switch region promoters after cytokine treatment (119, 120). We hypothesize that LPS induction of CSR might partially result from induction of YY1 binding to the 3'RR and  $E\mu$  enhancers leading to induced DNA loop formation.

Alternatively, YY1 may be controlled by stage-specific post-translational modifications, or by stage-specific interaction with other proteins. A number of YY1 post-translational modifications can regulate YY1 DNA binding (phosphorylation of serines 180 and 184, and threonines 348 and 378) (130–132), and YY1 is sumoylated on lysine 288 (133), which can control protein–protein interactions. Phosphorylation of serines 180 and 184 is mediated by Aurora B kinase and expression of this kinase peaks in splenic germinal center B cells (www.immgen.org) when CSR is active. Several studies demonstrated that YY1 subcellular localization is regulated during cell cycle progression and development (132, 134–137) suggesting that YY1 might also regulate subcellular localization of interacting partner proteins. In addition, apoptotic stimuli promote rapid translocation of YY1 from the cytoplasm to the nucleus in asynchronous HeLa cells (138). Thus, YY1 might function to increase transport of proteins from the cytoplasm to the nucleus via the nuclear pore.

During B cell development, YY1 expression levels remain relatively constant, as defined by transcript levels (www.immgen.org). However, YY1 protein levels are regulated in some systems yielding biological responses. This is most well studied in skeletal muscle



**FIGURE 5 | Model of YY1 recruitment of proteins to DNA needed for long-distance DNA interactions.** Using the IgH locus as a model, YY1 binding sites are indicated. Binding by YY1 then results in recruitment of condensin, cohesin, and PcG complex proteins. These proteins may form homotypic or heterotypic interactions to mediate long-distance DNA

interactions. Positions of the E<sub>μ</sub> enhancer, 3'RR enhancer, and various V genes are shown by arrows. Black rectangles represent various C<sub>H</sub> constant regions. YY1 also physically interacts with AID, and AID is able to interact with condensin, and cohesin complexes, and thus may contribute to DNA loop formation in germinal center B cells.

differentiation systems where YY1 expression levels drop as a result of proteolysis (24), and in cardiac disease conditions (139, 140). Thus, regulation of YY1 protein stability may control DNA loop formation.

It should be noted that RNA expression profiles of PcG proteins EZH2 and YAF2, as well as cohesin, and condensin subunit proteins SMC4, SMC2, SMC1, SMC3, CAP-G, CAP-H (BRRN1), and CAP-D2 all peak during B cell development at the pre-B cell stage (www.immgen.org). Expression levels are also high in pro-B cells, but peak in pre-B cells, then drop in immature B cell stages. This expression pattern is coincident with the timing of Ig rearrangement and is consistent with a role in Ig locus contraction and rearrangement. However, this timing is also coincident with high levels of proliferation in pre-B cells suggesting a possible effect of YY1 on the pre-B proliferative burst during development. All factors peak again in germinal center B cells (www.immgen.org) suggesting possible roles in proliferation, CSR, or somatic hypermutation.

Whatever the mode of locus accessibility or YY1 DNA binding, YY1 may then recruit proteins to DNA that are required for long-distance DNA interactions. As presented above, YY1 physically interacts with PcG, condensin, cohesin, and lamin proteins, all

involved in long-distance DNA interactions, and we have noted co-localization of some of these proteins with YY1 at the Igk locus (Figure 4). PcG proteins can mediate long-distance DNA interactions (60), and since YY1 recruits PcG proteins to DNA via the REPO domain (50, 61), we predict that this interaction will be important for long-distance interactions leading to DNA loop formation. Notably, condensin and cohesin complex proteins (105, 107), and lamin proteins (100–103) are all involved in long-distance DNA structures, suggesting that the DNA binding capacity of YY1 at IgH and Igk sequences may nucleate protein–protein interactions that govern DNA looping mechanisms. In addition to co-localization of YY1 and condensin proteins at the Igk locus, YY1 co-localizes with cohesin at the hs5–7 sites in the 3'RR enhancer (78, 81).

## MODELS OF YY1-MEDIATED LONG-DISTANCE DNA INTERACTIONS

Based upon: (a) the crucial nature of the YY1 REPO domain for B cell development, (b) the ability of this domain to recruit PcG proteins to DNA, (c) the physical interaction of the REPO domain with PcG, condensin, cohesin, and lamin proteins, (d) the co-localization of YY1, EZH2, and condensin proteins across the Igk

locus, (e) the co-localization of YY1 and cohesin proteins at the IgH 3'RR enhancer, (f) the effect of cohesin knock-down on CSR, (g) the effect of condensin subunit knock-down on V $\kappa$ -J $\kappa$  rearrangement, (h) the high levels of EZH2, YAF2, cohesin, and condensin proteins in pro-B, pre-B, and germinal center cells, (i) the critical role of YY1 in long-distance DNA loops in the IgH V region and 3' region, and (j) the regulatory role of YY1 in CSR, we propose the following mechanism. We propose that YY1 binds to sites spanning the IgH and Ig $\kappa$  loci. Concomitant with YY1 DNA binding, increased EZH2, YAF2, cohesin, and condensin subunit expression results in these proteins binding to the same DNA regions, presumably due to interactions with YY1. The nucleated PcG, cohesin, and condensin proteins then mediate long-distance interactions between the YY1 binding sites resulting in contraction of the Ig loci in looped or rosette structures (Figure 5). These loops then control somatic rearrangement of IgH and Ig $\kappa$  genes as well as CSR. Immediately upon maturation to the immature B cell stage, or upon maturation to plasma cells, EZH2, YAF2, cohesin, and condensin protein expression drops dramatically (www.immgen.org), thus facilitating de-contraction of the Ig loci, perhaps assisting in regulation of the allelic exclusion process, and causing a decrease in the inducible loops needed for CSR. In the case of CSR, it is intriguing that AID binds to many of the same factors that bind to YY1 (condensin, cohesin, and INO80 complexes) (123). Thus, YY1-AID physical interaction may also contribute to DNA loop formation (Figure 5).

Finally, it has been proposed that YY1 function in long-distance DNA interactions relates to the regulation of non-coding antisense transcripts in the IgH V<sub>H</sub> PAIR sequences (88). YY1 knock-out ablates some of these transcripts, and these transcripts have been proposed to play a role in IgH locus contraction (87, 88). Some RNA transcripts are known to regulate long-distance DNA interactions via interactions with the mediator complex (141). Whether YY1 functions in this mechanism is presently unclear.

## FUTURE STUDIES AND REMAINING QUESTIONS

A number of outstanding questions remain. (1) Is recruitment to DNA of proteins involved in DNA loop formation dependent upon YY1 DNA binding? (2) What mechanisms enable YY1 to function at distinct loci at various developmental stages? (3) Is YY1 function controlled by post-translational modifications? (4) Is YY1 controlled by stage-specific protein interactions? (5) What functions and domains of YY1 are needed for DNA looping, V(D)J rearrangement, and CSR? (6) What are the biochemical mechanisms for Ig locus contraction and for DNA loop formation? These questions and others are important for immune function and control of gene expression. The ubiquitous nature of YY1 and its involvement in looping at multiple loci (78, 87, 128) suggests that paradigms learned in the Ig systems will be globally applicable to other long-distance DNA interactions.

## ACKNOWLEDGMENTS

We thank Parul Mehra for critical comments on the manuscript and Arindam Basu for critical comments and assistance with diagrams. This work was supported by grants from the National Institutes of Health to Michael L. Atchison (grant numbers AI-079002 and GM082841).

## REFERENCES

- Church GM, Ephrussi A, Gilbert W, Tonegawa S. Cell-type-specific contacts to immunoglobulin enhancers in nuclei. *Nature* (1985) **313**:798–801. doi:10.1038/313798a0
- Ephrussi A, Church GM, Tonegawa S, Gilbert W. B lineage-specific interactions of an immunoglobulin enhancer with cellular factors in vivo. *Science* (1985) **227**:134–40. doi:10.1126/science.3917574
- Weinberger J, Baltimore D, Sharp PA. Distinct factors bind to apparently homologous sequences in the immunoglobulin heavy-chain enhancer. *Nature* (1986) **322**:846–8. doi:10.1038/322846a0
- Sen R, Baltimore D. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* (1986) **46**:705–16. doi:10.1016/0092-8674(86)90346-6
- Park K, Atchison ML. Isolation of a candidate repressor/activator, NF-E1 (YY-1,  $\delta$ ), that binds to the immunoglobulin  $\kappa$  3' enhancer and the immunoglobulin heavy-chain  $\mu$ E1 site. *Proc Natl Acad Sci U S A* (1991) **88**:9804–8. doi:10.1073/pnas.88.21.9804
- Shi Y, Seto E, Chang L-S, Shenk T. Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein. *Cell* (1991) **67**:377–88. doi:10.1016/0092-8674(91)90189-6
- Hariharan N, Kelley DE, Perry RP.  $\delta$ , a transcription factor that binds to downstream elements in several polymerase II promoters, is a functionally diverse zinc finger protein. *Proc Natl Acad Sci U S A* (1991) **88**:9799–803. doi:10.1073/pnas.88.21.9799
- Flanagan JR, Becker KG, Ennist DL, Gleason SL, Driggers PH, Levi B-Z, et al. Cloning of a negative transcription factor that binds to the upstream conserved region of Moloney murine leukemia virus. *Mol Cell Biol* (1992) **12**:38–44.
- Bushmeyer S, Park K, Atchison ML. Characterization of functional domains within the multifunctional transcription factor, YY1. *J Biol Chem* (1995) **270**:30213–20. doi:10.1074/jbc.270.50.30213
- Bushmeyer SM, Atchison ML. Identification of YY1 sequences necessary for association with the nuclear matrix and for transcriptional repression functions. *J Cell Biochem* (1998) **68**:484–99. doi:10.1002/(SICI)1097-4644(19980315)68:4<484::AID-JCBB>3.3.CO;2-F
- Galvin KM, Shi Y. Multiple mechanisms of transcriptional repression by YY1. *Mol Cell Biol* (1997) **17**:3723–32.
- Lee J-S, Galvin KM, See RH, Eckner R, Livingston D, Moran E, et al. Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. *Genes Dev* (1995) **9**:1188–98. doi:10.1101/gad.9.10.1188
- Lee T-C, Zhang Y, Schwartz RJ. Bifunctional transcriptional properties of YY1 in regulating muscle actin and c-myc gene expression during myogenesis. *Oncogene* (1994) **9**:1047–52.
- Lewis BA, Tullis G, Seto E, Horikoshi N, Weinmann R, Shenk T. Adenovirus E1A protein interact with the cellular YY1 transcription factor. *J Virol* (1995) **69**:1628–36.
- Yang W-M, Inouye C, Zeng Y, Bearss D, Seto E. Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. *Proc Natl Acad Sci U S A* (1996) **93**:12845–50. doi:10.1073/pnas.93.23.12845
- Thomas MJ, Seto E. Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? *Gene* (1999) **236**:197–208. doi:10.1016/S0378-1119(99)00261-9
- Austen M, Luscher B, Luscher-Firzlaff JM. Characterization of the transcriptional regulator YY1. The bipartite transactivation domain is independent of interaction with the TATA box-binding protein, transcription factor IIB, TAFII55, or cAMP-responsive element-binding protein (CPB)-binding protein. *J Biol Chem* (1997) **272**:1709–17.
- Donohoe ME, Zhang X, McGinnis L, Biggers J, Li E, Shi Y. Targeted disruption of mouse Yin Yang 1 transcription factor results in peri-implantation lethality. *Mol Cell Biol* (1999) **19**:7237–44.
- Gregoire S, Karra R, Passer D, Deutsch MA, Krane M, Feistritz R, et al. Essential and unexpected role of Yin Yang 1 to promote mesodermal cardiac differentiation. *Circ Res* (2013) **112**:900–10. doi:10.1161/CIRCRESAHA.113.259259
- Lu L, Sun K, Chen X, Zhao Y, Wang L, Zhou L, et al. Genome-wide survey by ChIP-seq reveals YY1 regulation of lincRNAs in skeletal myogenesis. *EMBO J* (2013) **32**:2575–88. doi:10.1038/emboj.2013.182
- Petkova V, Romanowski MJ, Suljoadikusumo I, Rohne D, Kang P, Shenk T, et al. Interaction between YY1 and the retinoblastoma protein. *J Biol Chem* (2001) **276**:7932–6. doi:10.1074/jbc.M007411200

22. Schlisio S, Halperin T, Vidal M, Nevins JR. Interaction of YY1 with E2Fs, mediated by RYBP, provides a mechanism for specificity of E2F function. *EMBO J* (2002) **21**:5775–86. doi:10.1093/emboj/cdf577
23. Sucharov CC, Dockstader K, McKinsey TA. YY1 protects cardiac myocytes from pathologic hypertrophy by interacting with HDAC5. *Mol Biol Cell* (2008) **19**:4141–53. doi:10.1091/mbc.E07-12-1217
24. Walowitz JL, Bradley ME, Chen S, Lee T. Proteolytic regulation of the zinc finger transcription factor YY1, a repressor of muscle-restricted gene expression. *J Biol Chem* (1998) **273**:6656–61. doi:10.1074/jbc.273.12.6656
25. Erkelaand S, Valkhof M, Heijmans-Antonissen C, Delwel R, Valk PJ, Hermans MH, et al. The gene encoding transcriptional regulator Yin Yang 1 (YY1) is a myeloid transforming gene interfering with neutrophilic differentiation. *Blood* (2003) **101**:1111–7. doi:10.1182/blood-2002-04-1207
26. Gabellini D, Green MR, Tupier R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* (2002) **110**:339–48. doi:10.1016/S0092-8674(02)00826-7
27. Li J, Shen H, Himmel KL, Dupuy AJ, Largaespada DA, Nakamura T, et al. Leukaemia disease genes: large scale cloning and pathway predictions. *Nat Genet* (1999) **23**:348–53. doi:10.1038/15531
28. Green MR, Monti S, Dalla-Favera R, Pasqualucci L, Walsh NC, Schmidt-Suppran M, et al. Signatures of murine B-cell development implicate Yy1 as a regulator of the germinal center-specific program. *Proc Natl Acad Sci U S A* (2011) **108**:2873–8. doi:10.1073/pnas.1019537108
29. Duke JL, Liu M, Yaari G, Khalil AM, Tomayko MM, Shlomchik MJ, et al. Multiple transcription factor binding sites predict AID targeting in non-Ig genes. *J Immunol* (2013) **190**:3878–88. doi:10.4049/jimmunol.1202547
30. Gordon S, Akopyan G, Garban H, Bonavida B. Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene* (2006) **25**:1125–42. doi:10.1038/sj.onc.1209080
31. Sui G. The regulation of YY1 in tumorigenesis and its targeting potential in cancer therapy. *Mol Cell Pharmacol* (2009) **1**:157–76. doi:10.4255/mcpharmacol.09.20
32. Zaravinos A, Spandidos DA. Yin Yang 1 expression in human tumors. *Cell Cycle* (2010) **9**:512–22. doi:10.4161/cc.9.3.10588
33. Castellano G, Torrisi E, Ligresti G, Malaponte G, Militello L, Russo AE, et al. The involvement of the transcription factor Yin Yang 1 in cancer development and progression. *Cell Cycle* (2009) **8**:1367–72. doi:10.4161/cc.8.9.8314
34. Castellano G, Torrisi E, Ligresti G, Nicoletti F, Malaponte G, Travalì S, et al. Yin Yang 1 overexpression in diffuse large B-cell lymphoma is associated with B-cell transformation and tumor progression. *Cell Cycle* (2010) **9**:557–63. doi:10.4161/cc.9.3.10554
35. Donohoe ME, Zhang L-F, Xu N, Shi Y, Lee JT. Identification of a Ctfc cofactor, YY1, for the X chromosome binary switch. *Mol Cell* (2007) **25**:43–56. doi:10.1016/j.molcel.2006.11.017
36. Jeon Y, Lee JT. YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* (2011) **146**:119–33. doi:10.1016/j.cell.2011.06.026
37. Kim JD, Hinz AK, Choo JH, Stubbs L, Kim J. YY1 as a controlling factor for the Peg3 and Gnas imprinted domains. *Genomics* (2007) **89**:262–9. doi:10.1016/j.ygeno.2006.09.009
38. Coull JJ, He G, Rucker MC, Dervan PB, Margolis DM. Targeted derepression of the human immunodeficiency virus type 1 long terminal repeat by pyrrole-imidazole polyamines. *J Virol* (2002) **76**:12349–54. doi:10.1128/JVI.76.23.12349-12354.2002
39. Dong XP, Stubenrauch F, Beyer-Finkler E, Pfister H. Prevalence of deletions of YY1-binding sites in episomal HPV16 DNA from cervical cancers. *Int J Cancer* (1994) **58**:803–8. doi:10.1002/ijc.2910580609
40. He G, Margolis DM. Counterregulation of chromatin deacetylation and histone deacetylase occupancy at the integrated promoter of human immunodeficiency virus type 1 (HIV-1) by the HIV-1 repressor YY1 and HIV-1 activator Tat. *Mol Cell Biol* (2002) **22**:2965–73. doi:10.1128/MCB.22.9.2965-2973.2002
41. Margolis DM, Somasundaran M, Green MR. Human transcription factor YY1 represses human immunodeficiency virus type 1 transcription and virion production. *J Virol* (1994) **68**:905–10.
42. May M, Dong XP, Beyer-Finkler E, Stubenrauch F, Fuchs PG, Pfister H. The E6/E7 promoter of extrachromosomal HPV16 DNA in cervical cancers escapes from cellular repression by mutation of target sequences for YY1. *EMBO J* (1994) **13**:1460–6.
43. Moriuchi M, Moriuchi H. YY1 transcription factor down-regulates expression of CCR5, a major coreceptor for HIV-1. *J Biol Chem* (2003) **278**:13003–7. doi:10.1074/jbc.M204980200
44. Schmidt M, Kedzia W, Gozdziak J. Intratype HPV16 sequence variation within LCR of isolates from asymptomatic carriers of cervical cancers. *J Clin Virol* (2001) **23**:65–77. doi:10.1016/S1386-6532(01)00189-5
45. Tan SH, Baker CC, Stunkel W, Bernard HU. A transcriptional initiator overlaps with a conserved YY1 binding site in the long control region of human papillomavirus type16. *Virology* (2003) **305**:486–501. doi:10.1006/viro.2002.1779
46. Veress G, Murvai M, Szarka K, Juhasz A, Konya J, Gergely L. Transcriptoinal activity of human papillomavirus type 16 variants having deletions in the long control region. *Eur J Cancer* (2001) **37**:1946–52. doi:10.1016/S0959-8049(01)00222-2
47. Brown JL, Mucci D, Whiteley M, Dirksen M-L, Kassiss JA. The *Drosophila* Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. *Mol Cell* (1998) **1**:1057–64. doi:10.1016/S1097-2765(00)80106-9
48. Girtton JR, Jeon SH. Novel embryonic and adult homeotic phenotypes are produced by pleiohomeotic mutations in *Drosophila*. *Dev Biol* (1994) **161**:393–407. doi:10.1006/dbio.1994.1040
49. Atchison L, Ghias A, Wilkinson F, Bonini N, Atchison ML. The YY1 transcription factor functions as a PcG protein in vivo. *EMBO J* (2003) **22**:1347–58. doi:10.1093/emboj/cdg124
50. Srinivasan L, Atchison ML. YY1 DNA binding and PcG recruitment requires CtBP. *Genes Dev* (2004) **18**:2596–601. doi:10.1101/gad.1228204
51. Srinivasan L, Pan X, Atchison ML. Transient requirements of YY1 expression for PcG transcriptional rescue and phenotypic rescue. *J Cell Biochem* (2005) **96**:689–99. doi:10.1002/jcb.20562
52. Basu A, Wilkinson FH, Colavita K, Fennelly C, Atchison ML. YY1 DNA binding and interaction with YAF2 is essential for Polycomb recruitment. *Nucleic Acids Res* (2013) **41**:1709–17. doi:10.1093/nar/gkt1187
53. Caretti G, Di Padova M, Micales B, Lyons GE, Sartorelli V. The Polycomb Ezh2 methyltransferase regulates muscle gene expression and skeletal muscle differentiation. *Genes Dev* (2004) **18**:2627–38. doi:10.1101/gad.1241904
54. Sing A, Pannell D, Karaïskakis A, Sturgeon K, Djabali M, Ellis J, et al. A vertebrate Polycomb response element governs segmentation of the posterior hind-brain. *Cell* (2009) **138**:885–97. doi:10.1016/j.cell.2009.08.020
55. Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE. Variable requirements for DNA-binding proteins at polycomb-dependent repressive regions in the human HOX clusters. *Mol Cell Biol* (2013) **33**:3274–85. doi:10.1128/MCB.00275-13
56. Mendenhall EM, Koche RP, Truong T, Zhou VW, Issac B, Chi AS, et al. GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet* (2010) **6**:e1001244. doi:10.1371/journal.pgen.1001244
57. Okulski H, Druck B, Bhalarao S, Ringrose L. Quantitative analysis of polycomb response elements (PREs) at identical genomic locations distinguishes contributions of PRE sequence and genomic environment. *Epigenetics Chromatin* (2011) **4**:4. doi:10.1186/1756-8935-4-4
58. Ringrose L, Rehmsmeier M, Dura J-M, Paro R. Genome-wide prediction of Polycomb/Trithorax response elements in *Drosophila melanogaster*. *Dev Cell* (2003) **5**:759–71. doi:10.1016/S1534-5807(03)00337-X
59. Su I-H, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT, et al. Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat Immunol* (2003) **4**:124–31. doi:10.1038/ni876
60. Lanzuolo C, Roure V, Dekker J, Bantignies F, Orlando V. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat Cell Biol* (2007) **9**:1167–74. doi:10.1038/ncb1637
61. Wilkinson FH, Park K, Atchison ML. Polycomb recruitment to DNA in vivo by the YY1 REPO domain. *Proc Natl Acad Sci U S A* (2006) **103**:19296–301. doi:10.1073/pnas.0603564103
62. Garcia E, Marcos-Gutierrez C, del Mar Lorente M, Moreno JC, Vidal M. RYBP, a new repressor protein that interacts with components of the mammalian Polycomb complex, and with the transcription factor YY1. *EMBO J* (1999) **18**:3404–18. doi:10.1093/emboj/18.12.3404
63. Kalenik JL, Chen D, Bradley ME, Chen SJ, Lee TC. Yeast two-hybrid cloning of a novel zinc finger protein that interacts with the multifunctional transcription factor YY1. *Nucleic Acids Res* (1997) **25**:843–9. doi:10.1093/nar/25.4.843

64. Trojer P, Cao AR, Gao Z, Li Y, Zhang J, Xu X, et al. L3MBTL2 protein acts in concert with PcG protein-mediated monoubiquitination of H2A to establish a repressive chromatin structure. *Mol Cell* (2011) **42**:438–50. doi:10.1016/j.molcel.2011.04.004
65. Woo CJ, Kharchenko PV, Daher L, Park PJ, Kingston RE. A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* (2010) **140**:99–110. doi:10.1016/j.cell.2009.12.022
66. Wilkinson F, Pratt H, Atchison ML. PcG recruitment by the YY1 REPO domain can be mediated by Yaf2. *J Cell Biochem* (2010) **109**:478–86. doi:10.1002/jcb.22424
67. Cobb RM, Oestreich KJ, Osipovich OA, Oltz EM. Accessibility control of V(D)J recombination. *Adv Immunol* (2006) **91**:45–109. doi:10.1016/S0065-2776(06)91002-5
68. Fugmann SD, Lee AI, Shockett PE, Villy IJ, Schatz DG. The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Annu Rev Immunol* (2000) **18**:495–527. doi:10.1146/annurev.immunol.18.1.495
69. Jung D, Giallourakis C, Mostoslavsky R, Alt FW. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol* (2006) **24**:541–70. doi:10.1146/annurev.immunol.23.021704.115830
70. Subrahmanyam R, Sen R. RAGs' eye view of the immunoglobulin heavy chain locus. *Semin Immunol* (2010) **22**:337–45. doi:10.1016/j.smim.2010.08.003
71. Fuxa M, Skok J, Souabni A, Salvaggio G, Roldan E, Busslinger M. Pax5 induces V-to-DJ rearrangements and constrains contraction of the immunoglobulin heavy-chain gene. *Genes Dev* (2004) **18**:411–22. doi:10.1101/gad.291504
72. Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* (2002) **296**:158–62. doi:10.1126/science.1068768
73. Roldán E, Fuxa M, Chong W, Martinez D, Novatchkova M, Busslinger M, et al. Locus "decontraction" and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nat Immunol* (2005) **6**:31–41. doi:10.1038/ni1150
74. Skok JA, Brown KE, Azuara V, Caparros M-L, Baxter J, Takacs K, et al. Non equivalent nuclear location of immunoglobulin alleles in B lymphocytes. *Nat Immunol* (2001) **2**:848–54. doi:10.1038/ni0901-848
75. Hesslein DG, Pflugh DL, Chowdhury D, Bothwell AL, Sen R, Schatz DG. Pax5 is required for recombination of transcribed, acetylated, 5' IgH V gene segments. *Genes Dev* (2003) **17**:37–42. doi:10.1101/gad.1031403
76. Jhunjhunwala S, van Zelm MC, Peak MM, Murre C. Chromatin architecture and the generation of antigen receptor diversity. *Cell* (2009) **138**:435–48. doi:10.1016/j.cell.2009.07.016
77. Sayegh C, Jhunjhunwala S, Riblet R, Murre C. Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. *Genes Dev* (2005) **19**:322–7. doi:10.1101/gad.1254305
78. Guo C, Gerasimova T, Hao H, Ivanova I, Chakraborty T, Selimyan R, et al. Two forms of loops generate the chromatin conformation of the immunoglobulin heavy-chain gene locus. *Cell* (2011) **147**:332–43. doi:10.1016/j.cell.2011.08.049
79. Jhunjhunwala S, van Zelm MC, Peak MM, Cutchin S, Riblet R, van Dogen JJ, et al. The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. *Cell* (2008) **133**:265–79. doi:10.1016/j.cell.2008.03.024
80. Lucas JS, Bossen C, Murre C. Transcription and recombination factories: common features? *Curr Opin Cell Biol* (2010) **23**:1–7. doi:10.1016/j.cob.2010.11.007
81. Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, et al. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the IgH locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci U S A* (2011) **108**:9566–71. doi:10.1073/pnas.1019391108
82. Degner SC, Wong TP, Jankevicius G, Feeney AJ. Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J Immunol* (2009) **182**:44–8.
83. Ebert A, McManus S, Tagoh H, Medvedovic J, Slavagiotto G, Novatchkova M, et al. The distal VH gene cluster of the IgH locus contains distinct regulatory elements with Pax5 transcription factor-dependent activity in pro-B cells. *Immunity* (2011) **34**:175–87. doi:10.1016/j.immuni.2011.02.005
84. Liu H, Schmidt-Suprain M, Shi Y, Hobeika E, Barteneva N, Jumaa H, et al. Yin Yang 1 is a critical regulator of B-cell development. *Genes Dev* (2007) **21**:1179–89. doi:10.1101/gad.1529307
85. Pan X, Papasani M, Hao Y, Calamito M, Wei F, Quinn WJ, et al. YY1 controls Igκ repertoire and B-cell development, and localizes with condensin on the Igκ locus. *EMBO J* (2013) **32**:1168–82. doi:10.1038/emboj.2013.66
86. Ribeiro de Almeida C, Stadhouders R, de Bruijn MJ, Bergen IM, Thongjuea S, Lenhard B, et al. The DNA-binding protein CTCF limits proximal Vκ recombination and restricts κ enhancer interactions to the immunoglobulin κ light chain locus. *Immunity* (2011) **35**:501–13. doi:10.1016/j.immuni.2011.07.014
87. Medvedovic J, Ebert A, Tagoh H, Tamir IM, Schwickert TA, Novatchkova M, et al. Flexible long-range loops in the VH gene region of the IgH locus facilitate the generation of a diverse antibody repertoire. *Immunity* (2013) **39**:229–44. doi:10.1016/j.immuni.2013.08.011
88. Verma-Gaur J, Torkamani A, Schaffer L, Head SR, Schork NJ, Feeney AJ. Non-coding transcription within the IgH distal V(H) region at PAIR elements affects the 3D structure of the IgH locus in pro-B cells. *Proc Natl Acad Sci U S A* (2012) **109**:17004–9. doi:10.1073/pnas.1208398109
89. Calame K, Atchison M. YY1 helps bring loose ends together. *Genes Dev* (2007) **21**:1145–52. doi:10.1101/gad.1559007
90. Feeney AJ, Verma-Gaur J. CTCF-cohesin complex: architect of chromatin structure regulates V(D)J rearrangement. *Cell Res* (2012) **22**:280–2. doi:10.1038/cr.2011.188
91. Guo C, Yoon HS, Franklin A, Jain S, Ebert A, Cheng H-L, et al. CTCF-binding elements mediate control of V(D)J recombination. *Nature* (2011) **477**:424–30. doi:10.1038/nature10495
92. Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnoli T, Mira-Bontenbal H, Brown KW, et al. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. *Nature* (2011) **476**:467–71. doi:10.1038/nature10312
93. Reynaud D, Demarco IA, Reddy KL, Schjerve H, Bertolino E, Chen Z, et al. Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat Immunol* (2008) **9**:927–36. doi:10.1038/ni.1626
94. Gordon SJ, Saleque S, Birshtein BK. Yin Yang 1 is a lipopolysaccharide-inducible activator of the murine 3' IgH enhancer, hs3. *J Immunol* (2003) **170**:5549–57.
95. Kimura K, Cuvier O, Hirano T. Chromosome condensation by a human condensin complex in *Xenopus* egg extracts. *J Biol Chem* (2001) **276**:5417–20. doi:10.1074/jbc.C000873200
96. Lehmann AR. The role of SMC proteins in the responses to DNA damage. *DNA Repair* (2005) **4**:309–14. doi:10.1016/j.dnarep.2004.07.009
97. Savidou E, Cobbe N, Steffensen S, Cotterill S, Heck MM. *Drosophila* CAP-D2 is required for condensin complex stability and resolution of sister chromatids. *J Cell Sci* (2005) **118**:2529–43. doi:10.1242/jcs.02392
98. Schmiesing JA, Gregson HC, Zhou S, Yokomori K. A human condensin complex containing hCAP-C-hCAP-E and CNAPl, a homolog of *Xenopus* XCAP-D2, colocalizes with phosphorylated histone H3 during the early stage of mitotic chromosome condensation. *Mol Cell Biol* (2000) **20**:6996–7006. doi:10.1128/MCB.20.18.6996-7006.2000
99. Steffensen S, Coelho PA, Cobbe N, Vass S, Costa M, Hassan B, et al. A role for *Drosophila* SMC4 in the resolution of sister chromatids in mitosis. *Curr Biol* (2001) **11**:295–307. doi:10.1016/S0960-9822(01)00096-3
100. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* (2008) **453**:948–51. doi:10.1038/nature06947
101. Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SW, Solovei I, Brugman W, et al. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell* (2010) **38**:603–13. doi:10.1016/j.molcel.2010.03.016
102. Pickersgill H, Kalverda B, de Wit E, Talhout W, Fornerod M, van Steensel B. Characterization of the *Drosophila melanogaster* genome at the nuclear lamina. *Nat Genet* (2006) **38**:1005–14. doi:10.1038/ng1852
103. van Bommel JG, Pagie L, Braunschweig U, Brugman W, Meuleman W, Kerckhoven RM, et al. The insulator protein SU(HW) fine-tunes nuclear lamina interactions of the *Drosophila* genome. *PLoS One* (2010) **5**:e15013. doi:10.1371/journal.pone.0015013
104. Losada A, Hirano T. Dynamic molecular linkers of the genome: the first decade of SMC proteins. *Genes Dev* (2005) **19**:1269–87. doi:10.1101/gad.1320505
105. D'Ambrosio D, Schmidt CK, Katou Y, Kelly G, Itoh T, Shirahige K, et al. Identification of cis-acting sites of condensin loading onto budding yeast chromosomes. *Genes Dev* (2008) **22**:2215–27. doi:10.1101/gad.1675708
106. Haering CH. Forward: the many fascinating functions of SMC protein complexes. *Chromosome Res* (2009) **17**:127–9. doi:10.1007/s10577-008-9008-8
107. Haeusler RA, Pratt-Hyatt M, Good PD, Gipson TA, Engelke DR. Clustering of yeast tRNA genes is mediated by specific association of condensin



- with tRNA gene transcription complexes. *Genes Dev* (2008) **22**:2204–14. doi:10.1101/gad.1675908
108. Hudson DF, Marshall KM, Earnshaw WC. Condensin: architect of mitotic chromosomes. *Chromosome Res* (2009) **17**:131–44. doi:10.1007/s10577-008-9009-7
  109. Hudson DF, Ohta S, Freisinger T, MacIsaac F, Sennels L, Alves F, et al. Molecular and genetic analysis of condensin function in vertebrate cells. *Mol Biol Cell* (2008) **19**:3070–9. doi:10.1091/mbc.E08-01-0057
  110. Hudson DF, Vagnarelli P, Gassmann R, Earnshaw WC. Condensin is required for nonhistone protein assembly and structural integrity of vertebrate mitotic chromosomes. *Dev Cell* (2003) **5**:323–36. doi:10.1016/S1534-5807(03)00199-0
  111. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, et al. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* (2010) **467**:430–5. doi:10.1038/nature09380
  112. Phillips-Cremins JE, Sauria ME, Sanyal A, Gerasimova TI, Lajoie BR, Bell JS, et al. Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* (2013) **153**:1281–95. doi:10.1016/j.cell.2013.04.053
  113. Wei Z, Gao F, Kim S, Yang H, Lyu J, An W, et al. Klf4 organizes long-range chromosomal interactions with the Oct4 locus in reprogramming and pluripotency. *Cell Stem Cell* (2013) **12**:36–47. doi:10.1016/j.stem.2013.05.010
  114. Zhang H, Jiao W, Sun L, Fan J, Chen M, Wang H, et al. Intrachromosomal looping is required for activation of endogenous pluripotency genes during reprogramming. *Cell Stem Cell* (2013) **13**:30–5. doi:10.1016/j.stem.2013.05.012
  115. Gosling KM, Makaroff LE, Theodoratos A, Kim Y-H, Whittle B, Rui L, et al. A mutation in a chromosome condensin II subunit, kleisin  $\beta$ , specifically disrupts T cell development. *Proc Natl Acad Sci U S A* (2007) **104**:12445–50. doi:10.1073/pnas.0704870104
  116. Strick TR, Kawaguchi T, Hirano T. Real-time detection of single-molecule DNA compaction by condensin I. *Curr Biol* (2004) **14**:874–80. doi:10.1016/j.cub.2004.04.038
  117. Peled JU, Kuang FL, Iglesias-Ussel MD, Roa S, Kalis SL, Goodman MF, et al. The biochemistry of somatic hypermutation. *Annu Rev Immunol* (2008) **26**:481–511. doi:10.1146/annurev.immunol.26.021607.090236
  118. Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. *Annu Rev Immunol* (2008) **26**:261–92. doi:10.1146/annurev.immunol.26.021607.090248
  119. Kumar S, Wuertfel R, Achour I, Lajoie B, Sen R, Dekker J, et al. Flexible ordering of antibody class switch and V(D)J joining during B-cell ontogeny. *Genes Dev* (2013) **27**:2439–44. doi:10.1101/gad.227165.113
  120. Wuertfel R, Wang L, Grigera F, Manis J, Selsing E, Perlot T, et al. S-Synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity* (2007) **27**:711–22. doi:10.1016/j.immuni.2007.09.007
  121. Xu Z, Zan H, Pone EJ, Mai T, Casali P. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat Rev Immunol* (2012) **12**:517–31. doi:10.1038/nri3216
  122. Garrett FE, Emelyanov AV, Sepulveda MA, Flanagan P, Volpi S, Li F, et al. 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* (2005) **25**:1511–25. doi:10.1128/MCB.25.4.1511-1525.2005
  123. Thomas-Claudepierre AS, Schiavo E, Heyer V, Fournier M, Page A, Robert I, et al. The cohesin complex regulates immunoglobulin class switch recombination. *J Exp Med* (2013) **210**:2495–502. doi:10.1084/jem.20130166
  124. Enervald E, Du L, Visnes T, Björkman A, Lindgren E, Wincent J, et al. A regulatory role for the cohesin loader NIPBL in nonhomologous end joining during immunoglobulin class switch recombination. *J Exp Med* (2013) **210**:2503–13. doi:10.1084/jem.20130168
  125. Cai Y, Jin J, Yao T, Gottschalk AJ, Swanson SK, Wu S, et al. YY1 functions with INO80 to active transcription. *Nat Struct Mol Biol* (2007) **14**:872–4. doi:10.1038/nsmb1276
  126. Wu S, Shi Y, Mulligan P, Gay F, Landry J, Liu H, et al. A YY1-INO80 complex regulates genomic stability through homologous recombination-based repair. *Nat Struct Mol Biol* (2007) **14**:1665–72. doi:10.1038/nsmb1332
  127. Zaprazna K, Atchison ML. YY1 controls immunoglobulin class switch recombination and nuclear activation-induced deaminase levels. *Mol Cell Biol* (2012) **32**:1542–54. doi:10.1128/MCB.05989-11
  128. Hwang SS, Kim YU, Lee S, Jang SW, Kim MK, Koh BH, et al. Transcription factor YY1 is essential for regulation of the Th2 cytokine locus and for Th2 cell differentiation. *Proc Natl Acad Sci U S A* (2013) **110**:276–81. doi:10.1073/pnas.1214682110
  129. Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged V<sub>H</sub> gene segments. *Cell* (1985) **40**:271–81. doi:10.1016/0092-8674(85)90141-2
  130. Baumeister P, Luo S, Skarnes WC, Sui G, Seto E, Shi Y, et al. Endoplasmic reticulum stress induction of the Grp78/BiP promoter: activating mechanisms mediated by YY1 and its interactive chromatin modifiers. *Mol Cell Biol* (2005) **25**:4529–40. doi:10.1128/MCB.25.11.4529-4540.2005
  131. Kassardjian A, Rizkallah R, Riman S, Renfro SH, Alexander KE, Hurt MM. The transcription factor YY1 is a novel substrate for Aurora B kinase at G2/M transition of the cell cycle. *PLoS One* (2012) **7**(11):e50645. doi:10.1371/journal.pone.0050645
  132. Rizkallah R, Hurt MM. Regulation of the transcription factor YY1 in mitosis through phosphorylation of its DNA-binding domain. *Mol Cell Biol* (2009) **29**:4766–76. doi:10.1091/mbc.E09-04-0264
  133. Deng Z, Wan M, Sui G. PIASy-mediated sumoylation of Yin Yang 1 depends on their interaction but not the RING finger. *Mol Cell Biol* (2007) **27**:3780–92. doi:10.1128/MCB.01761-06
  134. Ficzyz C, Esakiw C, Meyer D, Marley KE, Hurt M, Ovseneck N. Expression, activity, and subcellular localization of the Yin Yang1 transcription factor in *Xenopus* oocytes and embryos. *J Biol Chem* (2001) **276**:22819–25. doi:10.1074/jbc.M011188200
  135. Ficzyz C, Ovseneck N. The Yin Yang 1 transcription factor associates with ribonucleoprotein (mRNP) complexes in the cytoplasm of *Xenopus* oocytes. *J Biol Chem* (2002) **277**:8382–7. doi:10.1074/jbc.M110304200
  136. Palko L, Bass HW, Beyrouthy MJ, Hurt MM. The Yin Yang-1 (YY1) protein undergoes a DNA-replication-associated switch in localization from the cytoplasm to the nucleus at the onset of S phase. *J Cell Sci* (2004) **117**:465–76. doi:10.1242/jcs.00870
  137. Yue R, Kang J, Zhao C, Hu W, Tang Y, Liu X, et al. Beta-arrestin1 regulates zebrafish hematopoiesis through binding to YY1 and relieving polycomb group repression. *Cell* (2009) **139**:535–46. doi:10.1016/j.cell.2009.08.038
  138. Krippner-Heidenreich A, Walsemann G, Beyrouthy MJ, Speckgens S, Kraft R, Thole H, et al. Caspase-dependent regulation and subcellular redistribution of the transcriptional modulator YY1 during apoptosis. *Mol Cell Biol* (2005) **25**:3704–14. doi:10.1128/MCB.25.9.3704-3714.2005
  139. Sucharov CC, Helmke SM, Langer SJ, Perryman MB, Bristow M, Leinwand L. The Ku protein complex interacts with YY1, is up-regulated in human heart failure, and represses alpha myosin heavy-chain gene expression. *Mol Cell Biol* (2004) **24**:8705–15. doi:10.1128/MCB.24.19.8705-8715.2004
  140. Sucharov CC, Mariner P, Long C, Bristow M, Leinwand L. Yin Yang 1 is increased in human heart failure and represses the activity of the human alpha-myosin heavy chain promoter. *J Biol Chem* (2003) **278**:31233–9. doi:10.1074/jbc.M301917200
  141. Lai F, Orom UA, Cesaroni M, Beringer M, Taatjes DJ, Blobel GA, et al. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* (2013) **494**:497–501. doi:10.1038/nature11884

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 December 2013; paper pending published: 14 January 2014; accepted: 27 January 2014; published online: 10 February 2014.

Citation: Atchison ML (2014) Function of YY1 in long-distance DNA interactions. *Front. Immunol.* 5:45. doi: 10.3389/fimmu.2014.00045

This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Atchison. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Balancing proliferation with Ig $\kappa$ recombination during B-lymphopoiesis

Keith M. Hamel, Malay Mandal, Sophiya Karki and Marcus R. Clark\*

Department of Medicine, Section of Rheumatology, Gwen Knapp Center for Lupus and Immunology Research, The University of Chicago, Chicago, IL, USA

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

Roberta Pelanda, National Jewish Health and University of Colorado, USA

Bonnie B. Blomberg, University of Miami Miller School of Medicine, USA

## \*Correspondence:

Marcus R. Clark, Department of Medicine, Section of Rheumatology, Gwen Knapp Center for Lupus and Immunology Research, The University of Chicago, 924 East 57th Street, Chicago, IL 60637, USA  
e-mail: mclark@medicine.bsd.uchicago.edu

The essential events of B-cell development are the stochastic and sequential rearrangement of immunoglobulin heavy (Ig $\mu$ ) and then light chain (Ig $\kappa$  followed by Ig $\lambda$ ) loci. The counterpoint to recombination is proliferation, which both maintains populations of pro-B cells undergoing Ig $\mu$  recombination and expands the pool of pre-B cells expressing the Ig $\mu$  protein available for subsequent Ig $\kappa$  recombination. Proliferation and recombination must be segregated into distinct and mutually exclusive developmental stages. Failure to do so risks aberrant gene translocation and leukemic transformation. Recent studies have demonstrated that proliferation and recombination are each affected by different and antagonistic receptors. The IL-7 receptor drives proliferation while the pre-B-cell antigen receptor, which contains Ig $\mu$  and surrogate light chain, enhances Ig $\kappa$  accessibility and recombination. Remarkably, the principal downstream proliferative effectors of the IL-7R, STAT5 and cyclin D3, directly repress Ig $\kappa$  accessibility through very divergent yet complementary mechanisms. Conversely, the pre-B-cell receptor represses cyclin D3 leading to cell cycle exit and enhanced Ig $\kappa$  accessibility. These studies reveal how cell fate decisions can be directed and reinforced at each developmental transition by single receptors. Furthermore, they identify novel mechanisms of Ig $\kappa$  repression that have implications for gene regulation in general.

**Keywords: B cells, lymphopoiesis, recombination, proliferation, epigenetics**

## INTRODUCTION

Development of a diverse repertoire of peripheral B cells is dependent on the appropriate and ordered progression of B-lymphopoiesis. This process occurs through discrete developmental stages driven by the sequential rearrangement and expression of genes encoding the immunoglobulin heavy (Ig $\mu$ ) and then light chains (Ig $\kappa$  or Ig $\lambda$ ). Successful expression of a functional Ig $\mu$  capable of pairing with surrogate light chain (SLC) components and Ig $\alpha$ /Ig $\beta$  to form the pre-B-cell receptor (pre-BCR) at the cell surface is associated with a proliferative burst that expands the pool of pre-B cells expressing Ig $\mu$  prior to cell cycle exit and the rearrangement of Ig $\kappa$ . Proliferation and recombination must remain mutually exclusive to maintain genomic integrity and prevent excessive cell death or oncogenesis through aberrant translocations. Recent work has begun to uncover the molecular mechanisms dictating these developmental stages. Of particular interest, is the integration and opposition of the IL-7R and pre-BCR signaling pathways along with the effect of downstream epigenetic modifications on Ig $\kappa$  loci rearrangement and early B-cell proliferation.

## B-CELL DEVELOPMENT

Interactions with bone marrow (BM) stromal cells induce the differentiation of common lymphoid progenitor cells (CLPs), capable of generating B and T cells, into multipotential precursor-progenitor (pre-pro) B cells (1, 2). At this stage, initial Ig $\mu$  rearrangements occur at diversity (D<sub>H</sub>) and joining (J<sub>H</sub>) gene segments (3). Pre-pro-B cells are not committed to the B-cell lineage as some developing T cells bear Ig $\mu$  D<sub>H</sub>J<sub>H</sub> rearrangements.

Within IL-7 rich niches of the BM, pre-pro-B cells commit to the B-cell lineage through differentiation into progenitor (pro)-B cells expressing CD19 (4–6). IL-7 provides critical proliferative and survival signals needed to maintain the pool of pro-B cells. The hallmark event of pro-B cells is the completion of Ig $\mu$  rearrangement with the addition of a variable (V<sub>H</sub>) region to the D<sub>H</sub>J<sub>H</sub> segment. This process of recombination is mediated by the semi-random induction of double-stranded DNA breaks by the recombinase activating gene (Rag)-1 and Rag-2 proteins at recombination signal sequences (RSS) followed by non-homologous end joining (NHEJ) (7). Rag-mediated recombination of the antigen receptor loci is an essential and defining feature of B- and T-lymphopoiesis. Genetic mutation of the Rag genes results in severe combined immunodeficiency (SCID) in humans and mice (8–10).

Progression to the pre-B-cell stage of development is marked by the expression of a functional Ig $\mu$ , due to in-frame rearrangement at one locus, which can pair with SLC components, VpreB and  $\lambda$ 5, to form the pre-BCR at the cell surface (11). Early events following the expression of the pre-BCR serve to expand in number B-cell populations that have successfully rearranged Ig $\mu$  (12). Not all Ig $\mu$  chains effectively pair with SLC and therefore the pre-BCR checkpoint shapes the repertoire of Ig $\mu$  chains selected into the small pre-B-cell pool (13). In mice deficient in SLC, cells that escape by rearranging immunoglobulin light chain are preferentially autoreactive (14). Furthermore, conferring defined self-reactivity rescues SLC deficiency (15). However, it is not clear if this means that the pre-BCR censors autoreactivity or if autoreactivity, and ligation by self-antigen, is required to complement SLC deficiency.

Following poly-clonal expansion, late (small) pre-B cells migrate away from proliferation-inducing IL-7 rich niches of the BM, exit cell cycle, and begin to rearrange Igk genes (6). Final pairing of translated Igμ and Igk form the antigen-specific BCR on immature B cells which are then subjected to the mechanisms of tolerance that diminish autoreactivity in the naïve repertoire. Although the necessity of the IL-7R and pre-BCR for B-lymphopoiesis has long been appreciated, recent work has begun to detail the molecular mechanisms and downstream interplay of these pathways that drive B-cell development.

### IL-7R AND PRO-B CELLS FATE

Signaling through the IL-7R, which is a heterodimer of the IL-7Rα chain and the common γ chain, is essential for proliferation and survival of pro- and pre-B cells. *In vitro* culture assays demonstrated that pro-B cells and not pre-B cells proliferate in response to IL-7 (4). Accordingly, IL-7Rα-deficient mice demonstrate a significant impairment in B-lymphopoiesis beginning at the pro-B-cell stage (16–18). However, IL-7-deficient mice display a less severe defect in pro-B-cell development suggesting the IL-7Rα chain may participate in an additional signaling complex that compensates for loss of IL-7-induced signaling (17). Nonetheless, although pairing of IL-7Rα with alternative complexes may provide some compensation to IL-7-induced signaling, it is clear that the downstream components of the IL-7R pathway determine the pro-B-cell fate.

Through pairing with Janus kinase (JAK) 3 and JAK1, the IL-7R, upon activation, recruits and activates signal transducer and activator of transcription (STAT) 5a and b (19). STAT5 is critical for the biological effects of the IL-7R. B-cell development in mice deficient in both STAT5a and b is blocked at the pro-B stage, similar to IL-7Rα-deficient mice (20). Accordingly, constitutive activation (CA) of STAT5 in mice mostly restores B-lymphopoiesis in the absence of IL-7R signaling, while in humans, CA-STAT5 gene mutations have been identified in patients with acute lymphoblastic leukemia (21–23). Activated STAT5 primarily drives proliferation by inducing expression of the gene encoding cyclin D3, *Ccnd3* (23, 24). Pairing of cyclin D family members with cyclin-dependent kinases 4 and 6 (CDK4/6) during G<sub>1</sub> activates retinoblastoma protein (Rb) family members and E2f transcription factors to induce upregulation of cell cycle genes and suppress cell cycle inhibitors p27<sup>Kip1</sup> and p21<sup>Cip1</sup> (25). Although both cyclin D2 and D3 are expressed during B-cell development, only cyclin D3 can be found in complexes with CDK4/6 in pro-B cells (26). Moreover, a defect in early B-cell development is found only in *Ccnd3*<sup>-/-</sup> mice, while *Ccnd2*<sup>-/-</sup> mice display a later defect in peripheral B-cell proliferation (24, 27, 28). In addition to proliferative signals, STAT5 maintains survival of developing B cells through induction of several pro-survival genes including Mcl1, Bcl2, and Pim1 (22, 29, 30). Therefore, IL-7R-mediated activation of STAT5 represents a critical event in the expansion and stability of early B cells populations.

Pro-B cells are both proliferating and rearranging Igμ genes (4). Recent studies have provided some insights into how these incompatible processes are segregated to distinct populations within the pro-B-cell pool (31, 32). For example, it has been demonstrated that the core machineries of recombination and proliferation are

antagonistic. The Rag proteins are expressed in G<sub>0</sub>/G<sub>1</sub> and are degraded in dividing cells at the transition from G<sub>1</sub> to S phase (33). Cyclin A/CDK2 complexes induce cell cycle entry and inhibit the accumulation of Rag-2, while several CDK inhibitors, including p21<sup>Cip1</sup>, p27<sup>Kip1</sup>, and p57<sup>Kip2</sup> induce Rag-2 expression (34). This is because the cyclin A/CDK2 complex phosphorylates threonine 490 of Rag-2 targeting it for degradation by Skp2 (35). Mutation of threonine 490 results in persistence of Ig recombination in proliferating cells and increases the prevalence of chromosomal translocations and lymphoid malignancies (36). Impaired NHEJ accompanied with defective DNA-damage-induced apoptosis also increases the occurrence of leukemogenesis. Mice with combined deficiencies of the pro-apoptotic protein p53 with either XRCC4 or Ku80, both members of the NHEJ machinery, develop IgH-Myc translocations that promote pro-B leukemia (37, 38). Therefore, separation of proliferation and recombination is crucial to the avoidance of excessive B cells' death or development of B-cell leukemia.

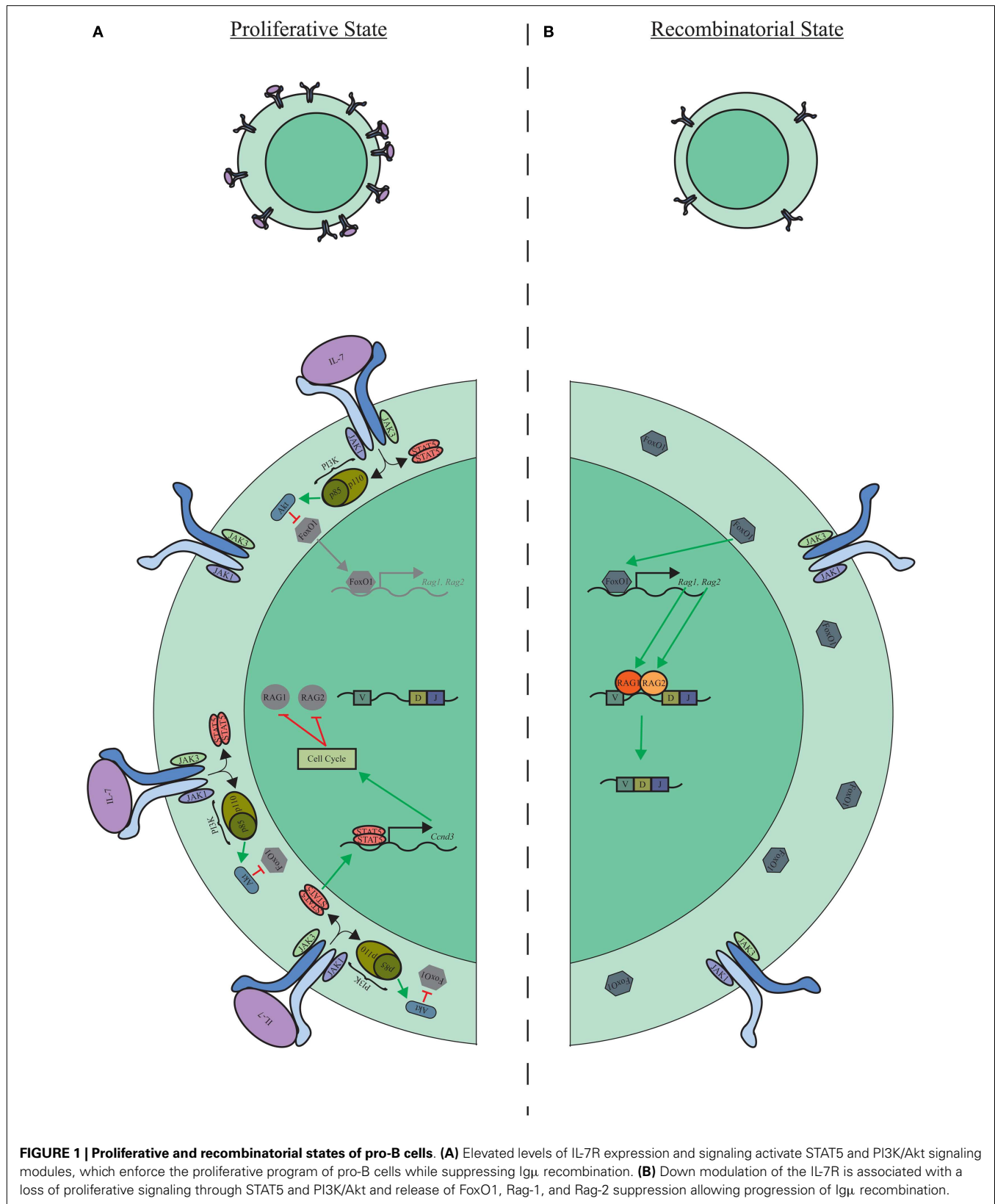
It is also now clear that the pro-B-cell compartment is not homogeneous but contains subpopulations of cells that express relatively high or low levels of the IL-7R. Furthermore, in these populations, IL-7R expression levels correlate with intracellular-activated STAT5 (39). These findings suggest a dynamic model where pro-B cells shift from proliferation to recombination through the oscillation of IL-7R expression (**Figure 1**). In contrast to oscillating between IL-7R high and low states, it is also possible that pro-B cells sequentially progress through IL-7R high and low stages. The mechanism driving IL-7R downregulation in pro-B cells, however, is still unknown. One possibility is through asymmetric cell division, where the accumulation of IL-7R toward IL-7-producing stromal cells results in distal daughter cells inheriting less IL-7R on their surface, therein, providing a transient decrease in STAT5 activation and the initiation of V<sub>H</sub>-D<sub>H</sub>J<sub>H</sub> rearrangement.

### PRE-BCR, PROLIFERATION, AND Igk REARRANGEMENT OF PRE-B CELLS

#### LARGE PRE-B CELLS

Cells transition to the pre-B-cell stage when Igμ pairs with SLC components, VpreB and λ5, along with the signaling module Igα/Igβ to form the pre-BCR at the cell surface. Initial expression of the pre-BCR is associated with a proliferative burst of early pre-B cells, also known as large pre-B cells, to expand the population of cells expressing a functional Igμ. Proper expression of the pre-BCR is critical to development as deficiencies of Igα, Igβ, or surface Igμ completely arrest B-lymphopoiesis while rearrangement and expression of Igk inefficiently rescues SLC deficiency (40–43). Activation of the pre-BCR requires the non-immunoglobulin domain of λ5, which mediates aggregation of the receptor (44–46). Although receptor aggregation is required, it is not clear if receptor aggregation is an intrinsic property of λ5 or if the SLC enables recognition of one or more selecting ligands within the BM (44, 47). Putative selecting ligands identified within the BM including heparin sulfate and galectin-1 have been suggested as natural ligands (48–50).

Concurrent to pre-BCR expression, large pre-B cells maintain IL-7R expression. It is within large pre-B cells that an additional



downstream target of IL-7R signaling important for B-cell development, the phosphoinositide 3-kinase (PI3K) pathway, plays a role (51, 52). The absence of PI3K has a definitive effect on peripheral B-cell proliferation, and selective deletion of the regulatory subunit p85 $\alpha$  or the combined catalytic subunits p110 $\alpha$  and p110 $\delta$  result in impairment of B-lymphopoiesis (53–55). However, the effects of PI3K on early B-cell proliferation appear to be within the initial proliferative events of pre-B cells, not pro-B cells. Deficiencies in p85 $\alpha$  or PTEN, a negative regulator of PI3K does not affect the number of pro-B cells in cycle, and the defect in development in p110 $\alpha$ - and p110 $\delta$ -deficient mice begins at the pre-B-cell stage (26, 52). Compared to cycling pro-B cells, large pre-B cells are indeed larger in size and display a heightened rate of proliferation (4). PI3K may be required in large pre-B cells to support increased protein synthesis and rapid cell division through increased glucose uptake and glycolytic activity by activated Akt, downstream of PI3K (56–58). Coincidentally, Akt is capable of enhancing survival by inhibiting pro-apoptotic pathways through direct repression of BAD and also indirectly by suppressing FoxO transcription factors, which induce Bim (59–62).

The pre-BCR is expressed on large pre-B cells and therefore has been thought to enhance proliferation in response to IL-7R signaling. Among, the signaling pathways common to the BCR and the IL-7R in the periphery, PI3K was an attractive candidate for any synergy that might occur between the two receptors. However, the pre-BCR does not efficiently couple to PI3K. Transfection of *Rag-2*<sup>-/-</sup> pro-B cells in the presence of IL-7 with a prearranged, functional *Ig $\mu$*  resulting in pre-BCR expression does not increase phospho-Akt activation and phospho-Akt levels are similar in pro and large pre-B cells (52). Furthermore, deletions of the genes encoding several pre-BCR downstream signaling components, including BLNK (SLP-65), Btk, and phospholipase  $\text{C}\gamma 2$  (PLC $\gamma 2$ ), result in a developmental block at the cycling pre-B-cell stage (63–65). Finally, re-expression of BLNK in deficient cells induces cell cycle arrest and Igk rearrangement (66). These observations indicate that the pre-BCR signals cell cycle exit rather than proliferation.

Therefore, the mechanisms driving the pre-B-cell proliferative burst remain unclear. It is possible that in pre-B cells, the pre-BCR has two signaling states, one pro-proliferative and one anti-proliferative (52, 67). However, the downstream effectors of such a pre-BCR-dependent proliferative pathway have yet to be identified. Alternatively, signaling mechanisms occurring independently of the pre-BCR could enhance IL-7R-mediated proliferation.

In addition to driving proliferation, signals through the IL-7R, and the downstream activation of STAT5, potentially repress Igk recombination (68). Activated STAT5 binds as a tetramer to a critical E-box-containing enhancer region of Igk, the intronic enhancer ( $\text{E}\kappa_i$ ), and tetrameric binding enables recruitment of the polycomb repressive complex (PRC2), which represses accessibility of the Igk region (69). Additionally, PI3K–Akt activation by the IL-7R represses recombination through indirect downregulation of Rag proteins (52). FoxO transcription factors induce Rag-1 and Rag-2 expression, however, repression of FoxO by the PI3K–Akt module inhibits Rag protein expression and inhibits recombination (70, 71). Therefore, beyond the intrinsic regulation of Rag proteins by the cell cycle machinery as described above, in large pre-B

cells, IL-7R signaling through STAT5, and the PI3K–Akt module, further enforce proliferation while suppressing pre-BCR-induced recombination.

### SMALL PRE-B CELLS

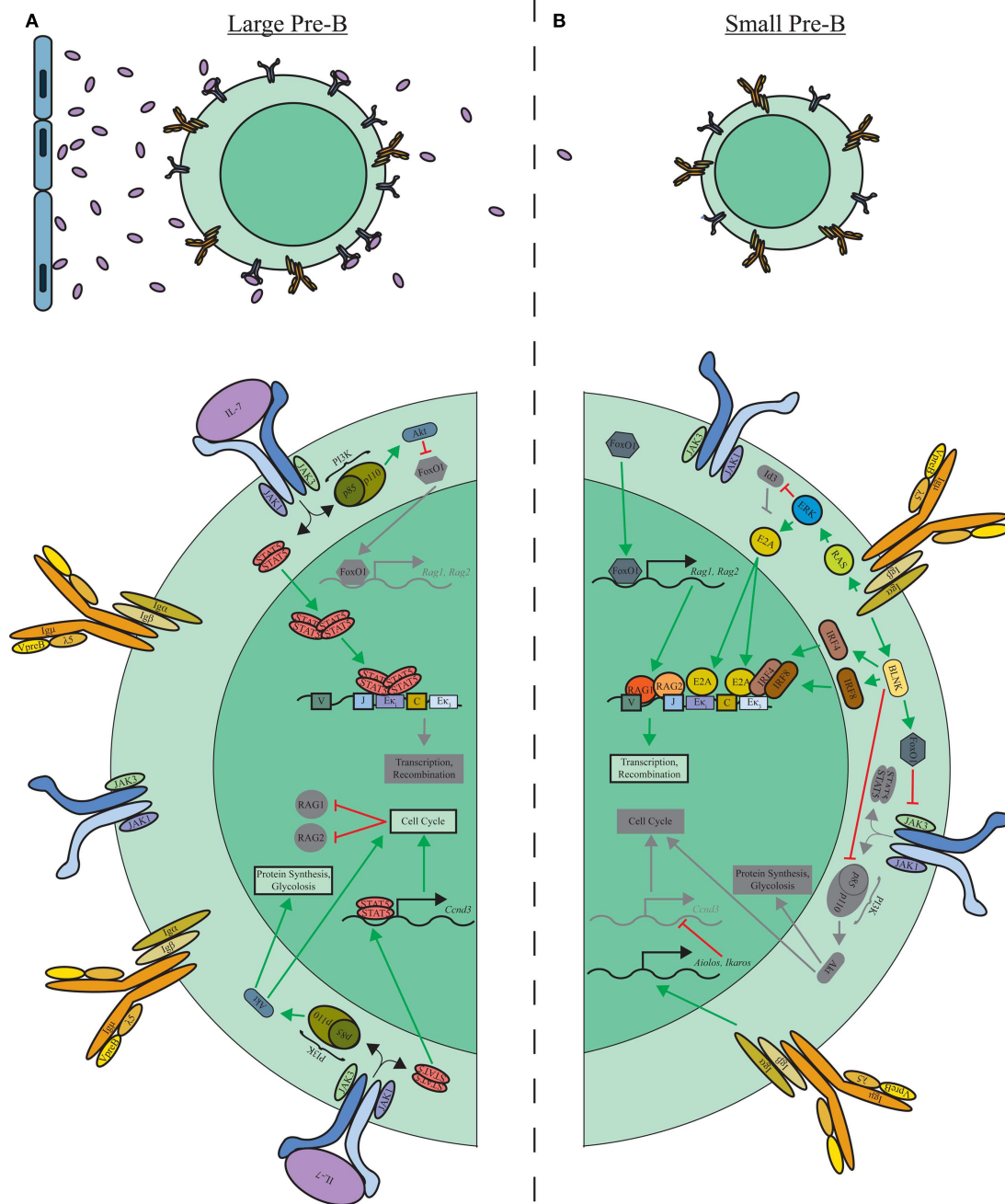
The transition from highly proliferative large pre-B cells to small resting pre-B cells undergoing Igk recombination is a pivotal point in normal B-lymphopoiesis. This transition is controlled by the signaling cascades downstream of the IL-7R and pre-BCR (Figure 2). As described below, the pre-BCR orchestrates Igk recombination, but cannot do so while the IL-7R is transmitting signals (23, 52, 68). First cells must escape IL-7 signaling, presumably through migration toward IL-7 low niches of the BM (6). Interestingly, upregulation of the interferon regulatory factor (IRF)-4 by the pre-BCR induces the expression of the chemokine receptor CXCR4 (68). The potential presence of the CXCR4 ligand, CXCL12, outside of IL-7 niches, may provide a mechanism by which early events of the pre-BCR enables movement into relatively IL-7-deficient niches and transition from proliferation-inducing signals (IL-7R) to those driving recombination (pre-BCR).

The opening of the Igk locus by the pre-BCR is predominately accomplished through activation of the Ras/Erk pathway (23, 72). Activated Erk induces E2A and inhibits the E2A repressor Id3 leading to an accumulation of free E2A within the nucleus (23, 73) that then binds the  $\text{E}\kappa_i$  and the Igk 3' enhancer ( $\text{E}\kappa_3$ ) (23). Escape from IL-7 signaling relieves tetrameric STAT5 occupancy of  $\text{E}\kappa_i$ , allowing E2A to bind, which promotes accessibility of the Igk loci for transcription and recombination (69). Genetic targeting of the E-boxes within  $\text{E}\kappa_i$  has demonstrated the importance of E2A recruitment in Igk recombination (74).

In addition to de-repressing Igk, loss of IL-7R signaling enhances specific pre-BCR-dependent and -independent mechanisms important for Igk recombination. Loss of IL-7R-induced PI3K–Akt activation results in increased FoxO expression. FoxO1 directly binds the Rag-1 and -2 genes and induces their expression (70). FoxO also binds and induces expression of the Syk and BLNK genes (52). The Syk/BLNK module induces the transcription factors IRF4 and 8, which bind the 3' Igk enhancer ( $\text{E}\kappa_3$ ) and enhance Igk accessibility (68, 75, 76). Furthermore, downstream of BLNK, activation of p38 MAP kinase further enhances FoxO activation thereby setting up a feed-forward loop that reinforces commitment to Igk recombination (52).

Pre-B-cell receptor signals additionally repress the proliferative program. FoxO1 represses surface expression of IL-7R in pre-B cells, while BLNK inhibits PI3K/Akt activation (52, 71). Pre-BCR signals also induce the expression of the transcription factors Aiolos and Ikaros (77, 78). These factors impede cell cycle by repression of Myc and cyclin D3 gene expression (23, 78). Accordingly, conditional deletion of Ikaros at the pro-B-cell stage of development results in a severe block in B-lymphopoiesis with an accumulation of cycling large pre-B cells (79). Ikaros might have a direct role in Igk recombination although the mechanisms remain to be defined (79). Collectively, downstream of the IL-7R and pre-BCR, these networks of feed-forward and feed-back mechanisms mediate the transition from proliferation to recombination and ensure sharp demarcation between each developmental state (80).





**FIGURE 2 | IL-7R and pre-BCR mediated transition of large pre-B to small pre-B cells. (A)** Localization of large pre-B cells near IL-7-producing stromal cells maintains IL-7-induced proliferation through *STAT5* and PI3K/Akt signaling modules. Additionally, tetrameric *STAT5* reinforces inhibition of Ig $\kappa$  recombination through direct binding to *E $\kappa$* . **(B)** Migration away from IL-7-rich niches limits IL-7R signaling allowing pre-BCR-induced

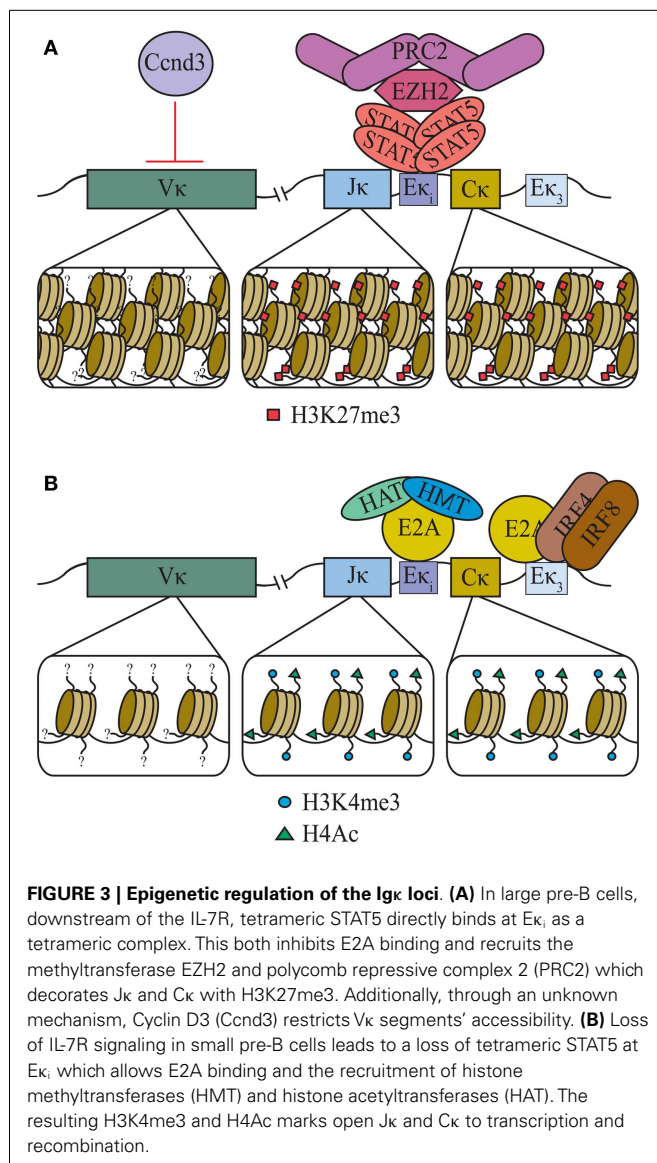
Ras/ERK and BLNK signaling modules to promote E2A and IRF4/IRF8 induction. Binding of these transcription factors to Ig $\kappa$  enhancer elements enables recombination in small pre-B cells. Additionally, the BLNK module, along with Aiolos and Ikaros, downstream of the pre-BCR inhibit proliferation by repressing IL-7R expression, PI3K/Akt activation, and *Ccnd3* transcription.

## EPIGENETIC REGULATION OF Ig $\kappa$ ACCESSIBILITY AND RECOMBINATION

### IL-7R AND PRE-BCR IMPOSED REGULATION OF Ig $\kappa$ ACCESSIBILITY

Chromatin structure and accessibility are fundamental to B-cell development. Recent evidence indicates that, at least in part,

accessibility of Ig genes is determined by post-translational epigenetic modifications of regional histone cores. Accessibility to recombination correlates with transcription (81) and indeed the primary effectors of epigenetic remodeling are transcription factors. It has become apparent that both *STAT5* and E2A regulate



Igk accessibility by determining the epigenetic landscape of the locus in pre-B cells (**Figure 3**). Initially, tetrameric STAT5, downstream of the IL-7R, recruits the histone methyltransferase Ezh2, which decorates the Igk locus with repressive histone 3 lysine 27 trimethylation (H3K27me3) marks (69). Following release from STAT5-mediated repression of Igk, E2A can access Eki, and marks the flanking Jk and Ck segments with activating H3K4 trimethylation (H3K4me3) and H4 acetylation (H4Ac) to promote an open chromatin structure (69, 82).

Interestingly, the above mechanisms of epigenetic regulation apply only to Jk and Ck and do not extend to the extensive Vκ regions (69). In fact, the Vκ regions are relatively devoid of any measured post-translational histone modifications identified for Ck and Jk [unpublished data and (83)]. Surprisingly, Vκ transcription is repressed by cyclin D3, through mechanisms that do not involve direct DNA binding (26). Instead, it appears that nuclear matrix-associated cyclin D3, and not that fraction associated

with CDK4/6, represses Vκ. The mechanisms by which cyclin D3 regulates Vκ transcription are not known, but might include controlling access to RNA polymerase II or nuclear positioning (84, 85). Regardless of mechanism, repression of Vκ accessibility by cyclin D3 provides a direct link between cell cycle transit and repression of Igk recombination.

### RAG-MEDIATED RECOMBINATION DEPENDS UPON EPIGENETIC MODIFICATIONS

Recombination events at Igk are also dependent on an open chromatin structure for accessibility of Rag proteins to RSS sites. RAG-mediated cleavage at RSS sites is restricted by a closed nucleosome structure (86–88). Histone modifications associated with open chromatin structures, including H3K4me3, histone 3 lysine 36 trimethylation (H3K36me3), H3Ac, and H4Ac correlate with recombination (89–91). Additionally, the recruitment of Rag-2 is dependent on the Rag-2 PHD domain binding to H3K4me3 (92, 93). The epigenetic regulation of JkCk, and the recruitment of RAG-2 to the marks of open chromatin, is consistent with current concepts that the JkCk region serves as the site of recombination (94). Furthermore, the JkCk region is anchored to the nuclear matrix and anchoring is necessary for efficient Igk recombination (95). This suggests that the recombination platform is relatively fixed and Vκ segments are recruited to it.

Although histone modifications at Jk and Ck have been associated with recombination and Rag-2 recruitment *in vivo*, there is no direct evidence that these modifications alone are capable of inducing RSS accessibility. In fact, *in vitro* experiments have demonstrated that hyperacetylation of histones is unable to overcome nucleosome-induced restriction of RSS sites and allow Rag-mediated recombination (87, 96). However, these extracellular *in vitro* experiments may lack additional lineage or stage-specific factors needed to translate epigenetic modifications into open chromatin. One such factor might be the SWI/SNF complex which can read specific epigenetic marks and open immunoglobulin gene loci for recombination (83, 97).

### CONCLUDING REMARKS

Recent observations have revealed that the IL-7R and the pre-BCR regulate complex networks of signaling and transcription cascades that direct and reinforce either pre-B-cell proliferation or Igk recombination. Central to understanding these networks is the clear demonstration that the IL-7R induces proliferation and represses Igk recombination and these biological activities are diametrically opposed by the pre-BCR. However, several questions still remain. For instance, if IL-7R signaling is constant in pro- and pre-B cells, and the pre-BCR does not provide a proliferative signal, what then is driving the large pre-B-cell proliferative burst? Additionally, although much effort has begun to describe how fate-determining transcription factors and epigenetic modifiers prime the required epigenetic landscape, little is known about the “readers” of these marks that impose and specify B-cell developmental events. The precise relationships between Igk transcription and recombination are unclear. Moreover, in the absence of epigenetic modifications, how is Vκ accessibility regulated? Further research into the molecular mechanisms that target and regulate the recombinatorial machinery to specific sites of the

Ig loci will be critical for understanding normal and pathogenic B-lymphopoiesis.

## ACKNOWLEDGMENTS

This work is supported by National Institute of Health (NIH)/National Institute of General Medical Sciences (NIGMS) grant numbers: 5R01GM088847, 5R01GM101090, and 5F32GM103143.

## REFERENCES

- Kondo M, Weissman IL, Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* (1997) **91**:661–72. doi:10.1016/S0092-8674(00)80453-5
- Izon D, Rudd K, Demuth W, Pear WS, Clendenin C, Lindsley RC, et al. A common pathway for dendritic cell and early B cell development. *J Immunol* (2001) **167**:1387–92.
- Jung D, Giallourakis C, Mostoslavsky R, Alt FW. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol* (2006) **24**:541–70. doi:10.1146/annurev.immunol.23.021704.115830
- Hardy RR, Carmack CE, Shinton SA, Kemp JD, Hayakawa K. Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J Exp Med* (1991) **173**:1213–25. doi:10.1084/jem.173.5.1213
- Schlissel MS. Regulating antigen-receptor gene assembly. *Nat Rev Immunol* (2003) **3**:890–9. doi:10.1038/nri1225
- Tokoyoda K, Egawa T, Sugiyama T, Choi BI, Nagasawa T. Cellular niches controlling B lymphocyte behavior within bone marrow during development. *Immunity* (2004) **20**:707–18. doi:10.1016/j.immuni.2004.05.001
- Gellert M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem* (2002) **71**:101–32. doi:10.1146/annurev.biochem.71.090501.150203
- Mombaerts P, Iacomini J, Johnson RS, Herrup K, Tonegawa S, Papaioannou VE. RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* (1992) **68**:869–77. doi:10.1016/0092-8674(92)90030-G
- Shinkai Y, Rathbun G, Lam KP, Oltz EM, Stewart V, Mendelsohn M, et al. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* (1992) **68**:855–67. doi:10.1016/0092-8674(92)90029-C
- de Miranda NF, Bjorkman A, Pan-Hammarstrom Q. DNA repair: the link between primary immunodeficiency and cancer. *Ann N Y Acad Sci* (2011) **1246**:50–63. doi:10.1111/j.1749-6632.2011.06322.x
- Nishimoto N, Kubagawa H, Ohno T, Gartland GL, Stankovic AK, Cooper MD. Normal pre-B cells express a receptor complex of mu heavy chains and surrogate light-chain proteins. *Proc Natl Acad Sci U S A* (1991) **88**:6284–8. doi:10.1073/pnas.88.14.6284
- Melchers F. The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. *Nat Rev Immunol* (2005) **5**:578–84. doi:10.1038/nri1649
- Melchers F, Ten Boekel E, Seidl T, Kong XC, Yamagami T, Onishi K, et al. Repertoire selection by pre-B-cell receptors and B-cell receptors, and genetic control of B-cell development from immature to mature B cells. *Immunol Rev* (2000) **175**:33–46. doi:10.1111/j.1600-065X.2000.imr017510.x
- Keenan RA, De Riva A, Corleis B, Hepburn L, Licence S, Winkler TH, et al. Censoring of autoreactive B cell development by the pre-B cell receptor. *Science* (2008) **321**:696–9. doi:10.1126/science.1157533
- Eschbach C, Bach MP, Fidler I, Pelanda R, Kohler F, Rajewsky K, et al. Efficient generation of B lymphocytes by recognition of self-antigens. *Eur J Immunol* (2011) **41**:2397–403. doi:10.1002/eji.201041344
- Peschon JJ, Morrissey PJ, Grabstein KH, Ramsdell FJ, Maraskovsky E, Gliniak BC, et al. Early lymphocyte expansion is severely impaired in interleukin 7 receptor-deficient mice. *J Exp Med* (1994) **180**:1955–60. doi:10.1084/jem.180.5.1955
- von Freeden-Jeffery U, Vieira P, Luciano L, McNeil T, Burdach SE, Murray R. Lymphopenia in interleukin (IL)-7 gene-deleted mice identifies IL-7 as a nonredundant cytokine. *J Exp Med* (1995) **181**:1519–26. doi:10.1084/jem.181.4.1519
- Corcoran AE, Riddell A, Krooshoop D, Venkitesan AR. Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. *Nature* (1998) **391**:904–7. doi:10.1038/36122
- O'Shea JJ, Plenge R. JAK and STAT signaling molecules in immunoregulation and immune-mediated disease. *Immunity* (2012) **36**:542–50. doi:10.1016/j.immuni.2012.03.014
- Yao Z, Cui Y, Watford WT, Bream JH, Yamaoka K, Hissong BD, et al. Stat5a/b are essential for normal lymphoid development and differentiation. *Proc Natl Acad Sci U S A* (2006) **103**:1000–5. doi:10.1073/pnas.0507350103
- Weber-Nordt RM, Egen C, Wehinger J, Ludwig W, Gouilleux-Gruart V, Mertelsmann R, et al. Constitutive activation of STAT proteins in primary lymphoid and myeloid leukemia cells and in Epstein-Barr virus (EBV)-related lymphoma cell lines. *Blood* (1996) **88**:809–16.
- Goetz CA, Harmon IR, O'Neil JJ, Burchill MA, Farrar MA. STAT5 activation underlies IL7 receptor-dependent B cell development. *J Immunol* (2004) **172**:4770–8.
- Mandal M, Powers SE, Ochiai K, Georgopoulos K, Kee BL, Singh H, et al. Ras orchestrates exit from the cell cycle and light-chain recombination during early B cell development. *Nat Immunol* (2009) **10**:1110–7. doi:10.1038/ni.1785
- Cooper AB, Sawai CM, Sicinska E, Powers SE, Sicinski P, Clark MR, et al. A unique function for cyclin D3 in early B cell development. *Nat Immunol* (2006) **7**:489–97. doi:10.1038/ni1324
- Sherr CJ, Roberts JM. CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev* (1999) **13**:1501–12. doi:10.1101/gad.13.12.1501
- Powers SE, Mandal M, Matsuda S, Miletic AV, Cato MH, Tanaka A, et al. Subnuclear cyclin D3 compartments and the coordinated regulation of proliferation and immunoglobulin variable gene repression. *J Exp Med* (2012) **209**:2199–213. doi:10.1084/jem.20120800
- Lam EW, Glassford J, Banerji L, Thomas NS, Sicinski P, Klaus GG. Cyclin D3 compensates for loss of cyclin D2 in mouse B-lymphocytes activated via the antigen receptor and CD40. *J Biol Chem* (2000) **275**:3479–84. doi:10.1074/jbc.275.5.3479
- Solvason N, Wu WW, Parry D, Mahony D, Lam EW, Glassford J, et al. Cyclin D2 is essential for BCR-mediated proliferation and CD5 B cell development. *Int Immunol* (2000) **12**:631–8. doi:10.1093/intimm/12.5.631
- Malin S, McManus S, Cobaleda C, Novatchkova M, Delogu A, Bouillet P, et al. Role of STAT5 in controlling cell survival and immunoglobulin gene recombination during pro-B cell development. *Nat Immunol* (2010) **11**:171–9. doi:10.1038/ni.1827
- Bednarski JJ, Nickless A, Bhattacharya D, Amin RH, Schlissel MS, Sleckman BP. RAG-induced DNA double-strand breaks signal through Pim2 to promote pre-B cell survival and limit proliferation. *J Exp Med* (2012) **209**:11–7. doi:10.1084/jem.20112078
- Bassing CH, Alt FW. The cellular response to general and programmed DNA double strand breaks. *DNA Repair* (2004) **3**:781–96. doi:10.1016/j.dnarep.2004.06.001
- Zhang Y, Gostissa M, Hildebrand DG, Becker MS, Boboila C, Chiarle R, et al. The role of mechanistic factors in promoting chromosomal translocations found in lymphoid and other cancers. *Adv Immunol* (2010) **106**:93–133. doi:10.1016/S0065-2776(10)06004-9
- Lin WC, Desiderio S. Cell cycle regulation of V(D)J recombination-activating protein RAG-2. *Proc Natl Acad Sci U S A* (1994) **91**:2733–7. doi:10.1073/pnas.91.7.2733
- Lee J, Desiderio S. Cyclin A/CDK2 regulates V(D)J recombination by coordinating RAG-2 accumulation and DNA repair. *Immunity* (1999) **11**:771–81. doi:10.1016/S1074-7613(00)80151-X
- Jiang H, Chang FC, Ross AE, Lee J, Nakayama K, Desiderio S. Ubiquitylation of RAG-2 by Skp2-SCF links destruction of the V(D)J recombinase to the cell cycle. *Mol Cell* (2005) **18**:699–709. doi:10.1016/j.molcel.2005.05.011
- Zhang L, Reynolds TL, Shan X, Desiderio S. Coupling of V(D)J recombination to the cell cycle suppresses genomic instability and lymphoid tumorigenesis. *Immunity* (2011) **34**:163–74. doi:10.1016/j.immuni.2011.02.003
- Difilippantonio MJ, Zhu J, Chen HT, Meffre E, Nussenzweig MC, Max EE, et al. DNA repair protein Ku80 suppresses chromosomal aberrations and malignant transformation. *Nature* (2000) **404**:510–4. doi:10.1038/35006670
- Gao Y, Ferguson DO, Xie W, Manis JP, Sekiguchi J, Frank KM, et al. Interplay of p53 and DNA-repair protein XRCC4 in tumorigenesis, genomic stability and development. *Nature* (2000) **404**:897–900. doi:10.1038/35009138
- Johnson K, Chaumeil J, Micsinai M, Wang JM, Ramsey LB, Baracho GV, et al. IL-7 functionally segregates the pro-B cell stage by regulating transcription of recombination mediators across cell cycle. *J Immunol* (2012) **188**:6084–92. doi:10.4049/jimmunol.1200368
- Kitamura D, Roes J, Kuhn R, Rajewsky K. A B cell-deficient mouse by targeted disruption of the membrane exon of the immunoglobulin mu chain gene. *Nature* (1991) **350**:423–6. doi:10.1038/350423a0

41. Gong S, Nussenzweig MC. Regulation of an early developmental checkpoint in the B cell pathway by Ig beta. *Science* (1996) **272**:411–4. doi:10.1126/science.272.5260.411
42. Pelandra R, Braun U, Hobeika E, Nussenzweig MC, Reth M. B cell progenitors are arrested in maturation but have intact VDJ recombination in the absence of Ig-alpha and Ig-beta. *J Immunol* (2002) **169**:865–72.
43. Shimizu T, Mundt C, Licence S, Melchers F, Martensson IL. VpreB1/VpreB2/lambda 5 triple-deficient mice show impaired B cell development but functional allelic exclusion of the IgH locus. *J Immunol* (2002) **168**:6286–93.
44. Ohnishi K, Melchers F. The nonimmunoglobulin portion of lambda5 mediates cell-autonomous pre-B cell receptor signaling. *Nat Immunol* (2003) **4**:849–56. doi:10.1038/ni959
45. Bankovich AJ, Raunser S, Juo ZS, Walz T, Davis MM, Garcia KC. Structural insight into pre-B cell receptor function. *Science* (2007) **316**:291–4. doi:10.1126/science.1139412
46. Vettermann C, Herrmann K, Albert C, Roth E, Bosl MR, Jack HM. A unique role for the lambda5 nonimmunoglobulin tail in early B lymphocyte development. *J Immunol* (2008) **181**:3232–42.
47. Kohler F, Hug E, Eschbach C, Meixlsperger S, Hobeika E, Kofer J, et al. Autoreactive B cell receptors mimic autonomous pre-B cell receptor signaling and induce proliferation of early B cells. *Immunity* (2008) **29**:912–21. doi:10.1016/j.immuni.2008.10.013
48. Gauthier L, Rossi B, Roux F, Termine E, Schiff C. Galectin-1 is a stromal cell ligand of the pre-B cell receptor (BCR) implicated in synapse formation between pre-B and stromal cells and in pre-BCR triggering. *Proc Natl Acad Sci U S A* (2002) **99**:13014–9. doi:10.1073/pnas.202323999
49. Bradl H, Wittmann J, Milius D, Vettermann C, Jack HM. Interaction of murine precursor B cell receptor with stroma cells is controlled by the unique tail of lambda 5 and stroma cell-associated heparan sulfate. *J Immunol* (2003) **171**:2338–48.
50. Mourcin F, Breton C, Tellier J, Narang P, Chasson L, Jorquera A, et al. Galectin-1-expressing stromal cells constitute a specific niche for pre-BII cell development in mouse bone marrow. *Blood* (2011) **117**:6552–61. doi:10.1182/blood-2010-12-323113
51. Corcoran AE, Smart FM, Cowling RJ, Crompton T, Owen MJ, Venkataraman AR. The interleukin-7 receptor alpha chain transmits distinct signals for proliferation and differentiation during B lymphopoiesis. *EMBO J* (1996) **15**:1924–32.
52. Ochiai K, Maienschein-Cline M, Mandal M, Triggs JR, Bertolino E, Sciammas R, et al. A self-reinforcing regulatory network triggered by limiting IL-7 activates pre-BCR signaling and differentiation. *Nat Immunol* (2012) **13**:300–7. doi:10.1038/ni.2210
53. Fruman DA, Snapper SB, Yballe CM, Davidson L, Yu JY, Alt FW, et al. Impaired B cell development and proliferation in absence of phosphoinositide 3-kinase p85alpha. *Science* (1999) **283**:393–7. doi:10.1126/science.283.5400.393
54. Suzuki H, Terauchi Y, Fujiwara M, Aizawa S, Yazaki Y, Kadowaki T, et al. Xid-like immunodeficiency in mice with disruption of the p85alpha subunit of phosphoinositide 3-kinase. *Science* (1999) **283**:390–2. doi:10.1126/science.283.5400.390
55. Okkenhaug K, Bilancio A, Farjot G, Priddle H, Sancho S, Peskett E, et al. Impaired B and T cell antigen receptor signaling in p110delta PI 3-kinase mutant mice. *Science* (2002) **297**:1031–4. doi:10.1126/science.1073560
56. Gottlob K, Majewski N, Kennedy S, Kandel E, Robey RB, Hay N. Inhibition of early apoptotic events by Akt/PKB is dependent on the first committed step of glycolysis and mitochondrial hexokinase. *Genes Dev* (2001) **15**:1406–18. doi:10.1101/gad.889901
57. Rathmell JC, Fox CJ, Plas DR, Hammerman PS, Cinalli RM, Thompson CB. Akt-directed glucose metabolism can prevent Bax conformation change and promote growth factor-independent survival. *Mol Cell Biol* (2003) **23**:7315–28. doi:10.1128/MCB.23.20.7315-7328.2003
58. Doughty CA, Bleiman BE, Wagner DJ, Dufort FJ, Mataraza JM, Roberts MF, et al. Antigen receptor-mediated changes in glucose metabolism in B lymphocytes: role of phosphatidylinositol 3-kinase signaling in the glycolytic control of growth. *Blood* (2006) **107**:4458–65. doi:10.1182/blood-2005-12-4788
59. Datta K, Bellacosa A, Chan TO, Tschlis PN. Akt is a direct target of the phosphatidylinositol 3-kinase. Activation by growth factors, v-src and v-Ha-ras, in Sf9 and mammalian cells. *J Biol Chem* (1996) **271**:30835–9. doi:10.1074/jbc.271.48.30835
60. Essafi A, Fernandez De Mattos S, Hassen YA, Soeiro I, Mufti GJ, Thomas NS, et al. Direct transcriptional regulation of Bim by FoxO3a mediates STI571-induced apoptosis in Bcr-Abl-expressing cells. *Oncogene* (2005) **24**:2317–29. doi:10.1038/sj.onc.1208421
61. Greer EL, Brunet A. FOXO transcription factors at the interface between longevity and tumor suppression. *Oncogene* (2005) **24**:7410–25. doi:10.1038/sj.onc.1209086
62. Huntington ND, Labi V, Cumano A, Vieira P, Strasser A, Villunger A, et al. Loss of the pro-apoptotic BH3-only Bcl-2 family member Bim sustains B lymphopoiesis in the absence of IL-7. *Int Immunol* (2009) **21**:715–25. doi:10.1093/intimm/dxp043
63. Jumaa H, Wollscheid B, Mitterer M, Wienands J, Reth M, Nielsen PJ. Abnormal development and function of B lymphocytes in mice deficient for the signaling adaptor protein SLP-65. *Immunity* (1999) **11**:547–54. doi:10.1016/S1074-7613(00)80130-2
64. Bai L, Chen Y, He Y, Dai X, Lin X, Wen R, et al. Phospholipase Cgamma2 contributes to light-chain gene activation and receptor editing. *Mol Cell Biol* (2007) **27**:5957–67. doi:10.1128/MCB.02273-06
65. Xu S, Lee KG, Huo J, Kurosaki T, Lam KP. Combined deficiencies in Bruton tyrosine kinase and phospholipase Cgamma2 arrest B-cell development at a pre-BCR+ stage. *Blood* (2007) **109**:3377–84. doi:10.1182/blood-2006-07-036418
66. Flemming A, Brummer T, Reth M, Jumaa H. The adaptor protein SLP-65 acts as a tumor suppressor that limits pre-B cell expansion. *Nat Immunol* (2003) **4**:38–43. doi:10.1038/ni862
67. Sandoval GJ, Graham DB, Bhattacharya D, Sleckman BP, Xavier RJ, Swat W. Cutting edge: cell-autonomous control of IL-7 response revealed in a novel stage of precursor B cells. *J Immunol* (2013) **190**:2485–9. doi:10.4049/jimmunol.1203208
68. Johnson K, Hashimshony T, Sawai CM, Pongubala JM, Skok JA, Aifantis I, et al. Regulation of immunoglobulin light-chain recombination by the transcription factor IRF-4 and the attenuation of interleukin-7 signaling. *Immunity* (2008) **28**:335–45. doi:10.1016/j.immuni.2007.12.019
69. Mandal M, Powers SE, Maienschein-Cline M, Bartom ET, Hamel KM, Kee BL, et al. Epigenetic repression of the Igk locus by STAT5-mediated recruitment of the histone methyltransferase Ezh2. *Nat Immunol* (2011) **12**:1212–20. doi:10.1038/ni.2136
70. Amin RH, Schlissel MS. Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. *Nat Immunol* (2008) **9**:613–22. doi:10.1038/ni.1612
71. Herzog S, Hug E, Meixlsperger S, Paik JH, Depinho RA, Reth M, et al. SLP-65 regulates immunoglobulin light chain gene recombination through the PI(3)K-PKB-Foxo pathway. *Nat Immunol* (2008) **9**:623–31. doi:10.1038/ni.1616
72. Shaw AC, Swat W, Davidson L, Alt FW. Induction of Ig light chain gene rearrangement in heavy chain-deficient B cells by activated Ras. *Proc Natl Acad Sci U S A* (1999) **96**:2239–43. doi:10.1073/pnas.96.5.2239
73. Kee BL, Quong MW, Murre C. E2A proteins: essential regulators at multiple stages of B-cell development. *Immunol Rev* (2000) **175**:138–49. doi:10.1111/j.1600-065X.2000.imr017514.x
74. Inlay MA, Tian H, Lin T, Xu Y. Important roles for E protein binding sites within the immunoglobulin kappa chain intronic enhancer in activating V kappa J kappa rearrangement. *J Exp Med* (2004) **200**:1205–11. doi:10.1084/jem.20041135
75. Lu R, Medina KL, Lancki DW, Singh H. IRF-4,8 orchestrate the pre-B-to-B transition in lymphocyte development. *Genes Dev* (2003) **17**:1703–8. doi:10.1101/gad.1104803
76. Ma S, Pathak S, Trinh L, Lu R. Interferon regulatory factors 4 and 8 induce the expression of Ikaros and Aiolos to down-regulate pre-B-cell receptor and promote cell-cycle withdrawal in pre-B-cell development. *Blood* (2008) **111**:1396–403. doi:10.1182/blood-2007-08-110106
77. Thompson EC, Cobb BS, Sabbatini P, Meixlsperger S, Parelho V, Liberg D, et al. Ikaros DNA-binding proteins as integral components of B cell developmental-stage-specific regulatory circuits. *Immunity* (2007) **26**:335–44. doi:10.1016/j.immuni.2007.02.010
78. Ma S, Pathak S, Mandal M, Trinh L, Clark MR, Lu R. Ikaros and Aiolos inhibit pre-B-cell proliferation by directly suppressing c-Myc expression. *Mol Cell Biol* (2010) **30**:4149–58. doi:10.1128/MCB.00224-10
79. Heizmann B, Kastner P, Chan S. Ikaros is absolutely required for pre-B cell differentiation by attenuating IL-7 signals. *J Exp Med* (2013) **210**:2823–32. doi:10.1084/jem.20131735
80. Clark MR, Mandal M, Ochiai K, Singh H. Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nat Rev Immunol* (2013) **14**:69–80. doi:10.1038/nri3570

81. Amin RH, Cado D, Nolla H, Huang D, Shinton SA, Zhou Y, et al. Biallelic, ubiquitous transcription from the distal germline Ig{kappa} locus promoter during B cell development. *Proc Natl Acad Sci U S A* (2009) **106**:522–7. doi:10.1073/pnas.0808895106
82. Beck K, Peak MM, Ota T, Nemazee D, Murre C. Distinct roles for E12 and E47 in B cell specification and the sequential rearrangement of immunoglobulin light chain loci. *J Exp Med* (2009) **206**:2271–84. doi:10.1084/jem.20090756
83. Du H, Ishii H, Pazin MJ, Sen R. Activation of 12/23-RSS-dependent RAG cleavage by hSWI/SNF complex in the absence of transcription. *Mol Cell* (2008) **31**:641–9. doi:10.1016/j.molcel.2008.08.012
84. Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* (2002) **296**:158–62. doi:10.1126/science.1068768
85. Fitzsimmons SP, Bernstein RM, Max EE, Skok JA, Shapiro MA. Dynamic changes in accessibility, nuclear positioning, recombination, and transcription at the Ig kappa locus. *J Immunol* (2007) **179**:5264–73.
86. Kwon J, Imbalzano AN, Matthews A, Oettinger MA. Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1. *Mol Cell* (1998) **2**:829–39. doi:10.1016/S1097-2765(00)80297-X
87. Golding A, Chandler S, Ballestar E, Wolffe AP, Schlissel MS. Nucleosome structure completely inhibits in vitro cleavage by the V(D)J recombinase. *EMBO J* (1999) **18**:3712–23. doi:10.1093/emboj/18.13.3712
88. Baumann M, Mamais A, McBlane F, Xiao H, Boyes J. Regulation of V(D)J recombination by nucleosome positioning at recombination signal sequences. *EMBO J* (2003) **22**:5197–207. doi:10.1093/emboj/cdg487
89. McMurry MT, Krangel MS. A role for histone acetylation in the developmental regulation of VDJ recombination. *Science* (2000) **287**:495–8. doi:10.1126/science.287.5452.495
90. Perkins EJ, Kee BL, Ramsden DA. Histone 3 lysine 4 methylation during the pre-B to immature B-cell transition. *Nucleic Acids Res* (2004) **32**:1942–7. doi:10.1093/nar/gkh523
91. Abarrategui I, Krangel MS. Noncoding transcription controls downstream promoters to regulate T-cell receptor alpha recombination. *EMBO J* (2007) **26**:4380–90. doi:10.1038/sj.emboj.7601866
92. Liu Y, Subrahmanyam R, Chakraborty T, Sen R, Desiderio S. A plant homeodomain in RAG-2 that binds hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity* (2007) **27**:561–71. doi:10.1016/j.immuni.2007.09.005
93. Matthews AG, Kuo AJ, Ramon-Maiques S, Han S, Champagne KS, Ivanov D, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* (2007) **450**:1106–10. doi:10.1038/nature06431
94. Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol* (2011) **11**:251–63. doi:10.1038/nri2941
95. Xu Y, Davidson L, Alt FW, Baltimore D. Deletion of the Ig kappa light chain intronic enhancer/matrix attachment region impairs but does not abolish V kappa J kappa rearrangement. *Immunity* (1996) **4**:377–85. doi:10.1016/S1074-7613(00)80251-4
96. McBlane F, Boyes J. Stimulation of V(D)J recombination by histone acetylation. *Curr Biol* (2000) **10**:483–6. doi:10.1016/S0960-9822(00)00449-8
97. Osipovich OA, Subrahmanyam R, Pierce S, Sen R, Oltz EM. Cutting edge: SWI/SNF mediates antisense Igh transcription and locus-wide accessibility in B cell precursors. *J Immunol* (2009) **183**:1509–13. doi:10.4049/jimmunol.0900896

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 January 2014; paper pending published: 28 February 2014; accepted: 19 March 2014; published online: 02 April 2014.

Citation: Hamel KM, Mandal M, Karki S and Clark MR (2014) Balancing proliferation with Igk recombination during B-lymphopoiesis. *Front. Immunol.* **5**:139. doi: 10.3389/fimmu.2014.00139

This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Hamel, Mandal, Karki and Clark. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# A novel Pax5-binding regulatory element in the Igk locus

Rena Levin-Klein, Andrei Kirillov, Chaggai Rosenbluh, Howard Cedar and Yehudit Bergman\*

Department of Developmental Biology and Cancer Research, Institute for Medical Research Israel-Canada, Hebrew University Medical School, Jerusalem, Israel

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

Michael Reth,  
Albert-Ludwigs-University, Germany  
James Hagman, National Jewish Health, USA  
Ranjan Sen, National Institutes of Health, USA

## \*Correspondence:

Yehudit Bergman, Department of Developmental Biology and Cancer Research, Institute for Medical Research Israel-Canada, Hebrew University Medical School, P.O.B. 12272, Ein Kerem, Jerusalem 91120, Israel  
e-mail: yehudit.bergman@huji.ac.il

The Igk locus undergoes a variety of different molecular processes during B cell development, including V(D)J rearrangement and somatic hypermutations (SHM), which are influenced by *cis* regulatory regions (RRs) within the locus. The Igk locus includes three characterized RRs termed the intronic (iEk), 3'Ek, and Ed enhancers. We had previously noted that a region of DNA upstream of the iEk and matrix attachment region (MAR) was necessary for demethylation of the locus in cell culture. In this study, we further characterized this region, which we have termed Dm, for demethylation element. Pre-rearranged Igk transgenes containing a deletion of the entire Dm region, or of a Pax5-binding site within the region, fail to undergo efficient CpG demethylation in mature B cells *in vivo*. Furthermore, we generated mice with a deletion of the full Dm region at the endogenous Igk locus. The most prominent phenotype of these mice is reduced SHM in germinal center B cells in Peyer's patches. In conclusion, we propose the Dm element as a novel Pax5-binding *cis* regulatory element, which works in concert with the known enhancers, and plays a role in Igk demethylation and SHM.

**Keywords:** V(D)J rearrangement, DNA methylation, B cell development, Pax5, somatic hypermutation

## INTRODUCTION

The B cell receptors (BCRs) are encoded in the mouse genome by the three immunoglobulin (Ig) loci, the IgH heavy chain locus, and the two light chain loci, Igk and Igl. In their germline conformations, the Ig loci do not give rise to functional proteins. It is only through a tightly regulated process of genome editing, termed V(D)J recombination, that the loci are reconfigured to allow transcription of an Ig gene in B cells. During the recombination process, the variable (V), diversity (D), and joining (J) segments are cleaved by the RAG complex and joined together into one continuous segment by the DNA repair machinery (1). Each rearrangement utilizes a single V, D (in the heavy chain), and J segment, and each B cell contains one productively rearranged heavy chain and light chain. In this way, B cells give rise to the multitude of antigen recognition specificities which constitutes the adaptive immune system.

The recombination of the different loci takes place in a developmentally staggered manner, with the IgH locus undergoing VDJ recombination first in the pro-B cell stage (2). Light chain rearrangement normally takes place only after a successful IgH rearrangement, which allows the cells to differentiate to the pre-B cell stage (3). In mice, the Igk locus is the primary source for the BCR light chain and will undergo preferential rearrangement. The recombination of the different loci is kept tightly separated, despite the fact that the enzymatic machinery responsible for the processes is essentially the same and is present at both the pro- and pre-B cell stages. The light chain loci are maintained in an inaccessible chromatin state *via* epigenetic mechanisms prior to the pre-B cell stage, at which point they become available to the rearrangement machinery (3, 4). One such epigenetic mark is DNA methylation, a mark that is established at the Igk locus during early embryonic development and which is hereditarily

maintained during cell division (5). DNA methylation has been shown to block the activity of the rearrangement machinery *in vitro* (6). The Igk locus undergoes selective demethylation at the pre-B cell stage, immediately prior to rearrangement (5, 7, 8). The rearranged Igk allele is unmethylated from that stage onward, while alleles which do not undergo rearrangement remain methylated, even at the mature B cell stage. The low level of methylation is significant for an additional stage of Igk editing during B cell development, namely for efficient somatic hypermutation (SHM), which will allow affinity maturation of the BCR in activated mature B cells (9). Methylated pre-rearranged Igk sequences do not undergo proper SHM at this stage, whereas identical unmethylated sequences do (10).

The stage-specific transcription, rearrangement, and chromatin structure of the Igk gene is mediated by regulatory sequences within and in proximity to the locus. The locus contains three characterized enhancers, including an intronic enhancer (iEk) (11), located in the intron between the Jk segments and the Ck exon and two enhancers situated a few 1000 bases downstream of the Ck exon, termed 3'Ek (12) and Ed (13). These enhancers work in cooperation to promote stage-specific chromatin accessibility, DNA demethylation, V to J rearrangement, heightened transcription of the locus, and SHM in activated B cells, with different enhancers contributing to a varying extent to each one of these processes. iEk and 3'Ek have been implicated in promoting accessibility and rearrangement of the locus in pre-B cells (14–16), while 3'Ek and Ed strongly effect the level of transcription and SHM in mature B cells (17, 18), neither of which is significantly affected by the deletion of iEk (14, 18). All of the three enhancers contribute together to the demethylation of the locus (16, 19). Replacement of iEk with the IgH intronic  $\mu$  enhancer is enough to change the rearrangement timing of the locus to the earlier pro-B cell stage,

showing that it is indeed these sequences which direct the temporal precision of the developmental program (20).

Other than the enhancers, there are a number of additional regulatory elements surrounding the Igk locus, increasing the complexity of the regulation. The recently discovered HS10 element, which lies downstream of Ed, appears to mostly function in plasma cells. While itself being a weak enhancer, HS10 acts as a co-enhancer to strengthen the activities of 3'Eκ and Ed (21). A matrix attachment region (MAR) lies immediately adjacent to iEκ and mediates connections between the locus and the nuclear matrix (22).

The activities of the *cis* regulatory elements are mediated by various transcription factors, which either activate or repress the enhancer activity. Many of these transcription factors are master regulators of the B cell lineage, which are important for maintaining B cell identity, such as E2A and PU.1 that bind sites in iEκ and 3'Eκ and substantially contribute to the enhancer activity (23–27). However, binding of Pax5, a master regulator of B cell identity, has been surprisingly missing from these enhancers in mature B cells. While binding sites have been identified in 3'Eκ (24, 25, 28), as well as in K-I and K-II (29, 30), which are regulatory regions (RR) upstream of the Jκ segments, Pax5 plays an inhibitory role in this context and is released during the pre-B cell stage when the locus is activated. This is despite the fact that Pax5 itself is necessary for the active induction of the locus (31).

In this work, we characterize a region adjacent to the MAR/iEκ elements. We had previously identified this element as a participant in the demethylation process of the Igk locus in cell culture and thereby designated it Dm (32). Here, we find that this element binds Pax5 in B cell stages from the pre-B cell stage and onward. It is necessary for demethylation of a pre-rearranged Igk transgene, but deletion of the element in the endogenous locus does not affect the demethylation process. We find that the element contributes to efficient SHM of the Igk locus, indicating that the Dm element functions at more than one stage of B cell development.

## MATERIALS AND METHODS

### MICE

Targeted mice were backcrossed for 10 generations on a BALB/c background. Igk<sup>ΔDm/ΔDm</sup> mice were bred with wild-type (WT) BALB/c to produce Igk<sup>WT/ΔDm</sup> mice. Human Cκ knock-in mice (33) (gift from M. Nussenzweig) were bred with either WT BALB/c or ΔDm BALB/c to produce Igk<sup>WT/WT</sup>Cκ<sup>h/m</sup> and Igk<sup>WT/ΔDm</sup>Cκ<sup>h/m</sup> mice, respectively. Igk<sup>WT/ΔDm</sup> mice were bred with CAST/EiJ (Cast) mice (Jackson Laboratory) to produce BALB/c/Cast Igk<sup>WT/WT</sup> and BALB/c/Cast Igk<sup>ΔDm/WT</sup> littermates. Rag1<sup>-/-</sup> mice (Jackson Laboratories) were bred onto a Cast background. Igk<sup>ΔDm/ΔDm</sup> were bred onto a C57BL/6 Rag1<sup>-/-</sup> (B6) background containing the 3H9 IgH chain transgene (IgH<sup>+</sup>). CAST/EiJ Rag1<sup>-/-</sup> mice were bred with C57BL/6 Rag1<sup>-/-</sup> IgH<sup>+</sup> either with or without a deletion of the Dm element, giving rise to B6/Cast Rag1<sup>-/-</sup> IgH<sup>+</sup> Igk<sup>ΔDm/WT</sup> and B6/Cast Rag1<sup>-/-</sup> IgH<sup>+</sup> Igk<sup>WT/WT</sup> mice, respectively. Mice were housed in specific pathogen-free conditions at the Hebrew University Medical School animal facility. Transgenic mouse lines Lκ, LκΔDm, and LκΔ70 were produced, using the constructs described in the Section “Targeting Constructs,” at the Hadassah Hospital Medical School Transgenic Unit.

Two independent founder lines were produced for the Lκ transgene, four for the LκΔDm and three for the LκΔ70. The copy number of the transgene for each founder line varied from low (two insertions) to high (20 insertions) with most lines having a moderate number of insertions (four to eight insertions). All animal procedures were approved by the Animal Care and Use Committee of the Hebrew University of Jerusalem.

### TARGETING CONSTRUCTS

The LκΔDm construct was prepared using the following steps; the 4.3-kb *KpnI*–*KpnI* fragment, containing VκJ5–Cκ sequence, was excised from the Lκ plasmid (34) and cloned into the *KpnI* site of the Bluescript vector which was modified to destroy the polylinker *XbaI* site. The resulting pBSKpn2 plasmid was cut at unique compatible *XbaI* and *NsiI* sites, and recirculized, resulting in the deletion of 930 bp *XbaI*–*NsiI* Dm fragment from the JκCκ intron. The *KpnI*–*KpnI* Dm-deleted fragment was excised and reinserted into the Lκ plasmid, resulting in the LκΔDm construct.

The LκΔ70 construct was prepared using the following steps; the *HindIII*-blunt *TaqI* 2.6-kb fragment, containing the germline Jκ region, was cloned to *HindIII*–*EcoRV* sites of the Bluescript vector. Next, a blunted *BstEII*–*BglII* 2-kb fragment containing the Cκ exon was cloned into blunted *EcoRI*–*BamHI* sites of the previously described Jκ containing Bluescript vector to yield the p-Δ70 construct, which had the 70 bp *TaqI*–*BstEII* deletion introduced into the *HindIII*/BglII 5.6-kb JκCκ germline sequence. The 1-kb intact intronic *XbaI*–*HindIII* region of pBSKpn2 plasmid (previously described, containing the *KpnI*–*KpnI* fragment from the Lκ plasmid) was replaced with *XbaI*–*HindIII* fragment bearing the 70 bp deletion, excised from p-Δ70. The 4.2-kb *KpnI*–*KpnI* fragment with the 70 bp deletion was excised from the resulting pBSKpn2 Δ70 plasmid and cloned back into the Lκ plasmid, replacing the original 4.3-kb *KpnI*–*KpnI* sequence, and yielding the LκΔ70 construct.

The ΔDm targeting vector was prepared as using the following steps; a short arm of homology (neo-SAH) plasmid was constructed by using *BanII* (ends filled with Klenow) and *NsiI* to excise the 1.25-kb MAR and Eik containing fragment from the pBκMAR plasmid. This fragment was cloned into the sticky *XbaI* and blunted *PstI* sites of the Bluescript vector. This construct was next cut at the polylinker sites *Clal* and *EcoRI* and used for insertion of the 1.26-kb *NotI*–*XbaI* loxP flanked neoR gene fragment from the pMMneoflox-8 plasmid (all restriction ends were made blunt by reaction with the Klenow fragment), a long arm of homology (TK-LAH) plasmid was constructed by excision of the 7.1-kb *PstI*–*PstI* germline Jκ–Cκ region containing fragment from pSPIg8 plasmid (ends were blunted by reaction with T4 polymerase) and ligation into the *HindIII* site (blunted with reaction with Klenow fragment) of pIC19R/MC1-TK. The final ΔDm targeting vector was produced by cloning of the 8.9-kb *XhoI*–*Sall* fragment from TK-LAH into the neo-SAH *Sall* polylinker site. Targeting strategy is illustrated in Figure S1 in Supplementary Material.

### CELLS AND CULTURES

All cells in this manuscript were grown in RPMI 1640 medium (Gibco) supplemented with 10% fetal calf serum, 2 mM

L-Glutamine, 100 µg/ml penicillin, 100 µg/ml streptomycin, and 50 µM 2-mercaptoethanol. BaF3 cell medium was additionally supplemented with IL3 secreted by WEHI-3b cells. IL-7-dependent pre-B cell cultures used for chromatin immunoprecipitation (ChIP) analysis were performed as has been previously described (35). COP8 cells were transiently transfected with a Pax5 expression plasmid (gift from M. Busslinger) using the DEAE dextran method (36).

## ISOLATION AND ANALYSIS OF LYMPHOID CELLS FROM BONE MARROW AND SPLEEN

Bone marrow cells from femur and tibia bones were flushed out with PBS using a 25 G syringe needle. Spleens were disrupted and pulp dispersed in PBS. Erythrocytes were lysed with RBC lysis solution (Biological industries) and cells were washed. When indicated, cells were isolated on magnetic MACS columns (Miltenyi Biotech) by positive selection with either αCD19 magnetic beads or streptavidin magnetic beads and biotinylated αB220 (Miltenyi Biotech), according to the manufacturer's instructions. Cell purity following isolation was assayed as <95% by flow cytometry (LSR II, BD Bioscience).

Cells from erythrocyte disrupted spleens and bone marrows were stained with the antibodies indicated and cellular composition was analyzed by flow cytometry (LSR II, BD Bioscience). The antibodies used in this report include anti-mouse-Igκ-PE (Southern Biotech), anti-human-Igκ-FITC (Southern Biotech), anti-IgM-APC (eBioscience), anti-B220-PerCP-Cy5.5 (Biolegend), anti-CD43-PE (Biolegend), anti-IgD-FITC (eBioscience). Flow cytometry output was analyzed using Flowing Software v2.5.0 (Turku Centre for Biotechnology).

## ANALYSIS OF DNA METHYLATION BY SOUTHERN HYBRIDIZATION

Cellular genomic DNA (5–15 µg) was digested with the specified enzymes, electrophoresed in native (Tris–acetate) agarose gels, denatured and transferred to nitrocellulose. DNA was then hybridized with the specific radioactive probes and analyzed by autoradiography (37). Hybridization was carried out at 65°C for 16 h. The degree of methylation was measured semiquantitatively using a PhosphorImager BAS-1800 (Fuji) and Tina2.10 g software (IsotopenMedgerate GmbH).

## NUCLEAR EXTRACT PREPARATION

Cells ( $3\text{--}5 \times 10^6$ ) were washed in PBS, resuspended in low salt buffer (10 mM HEPES pH 7.9, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, 0.5 mM PMSF, 20 µg/ml aprotinin, 10 µg/ml leupeptin) and incubated for 10 min on ice. NP-40 was then added to a final concentration of 0.66%, the mixture was vortexed briefly and centrifuged for 30 s, 16,000 g. Nuclei were resuspended in high salt buffer (20 mM HEPES pH 7.9, 0.4 mM KCl, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, 1 mM PMSF, 20 µg/ml aprotinin, 10 µg/ml leupeptin) and rotated for 20 min at 4°C. Nuclear debris was removed by centrifugation at 16,000 g for 20 min at 4°C.

## ELECTROPHORETIC MOBILITY SHIFT ASSAY

Oligonucleotide probes were end-labeled with α<sup>32</sup>P-dCTP using Klenow fragment. Two micrograms of nuclear extract was

incubated with 0.3 ng of the radioactive double strand probe in a solution containing 2 µg poly-dI-dC, 10 mM Tris–HCl pH 7.9, 10% glycerol, 100 mM KCl, and 4 mM DTT for 20 min at 25°C. In competition assay, 100-fold molar excess of an unlabeled probe was preincubated for 10 min prior to the addition of the radiolabeled probe. In supershift assays, the indicated antisera (antibodies A1 and A2 kindly provided by Meinrad Busslinger) were added to the nuclear extract 15 min prior to the addition of the probe. Samples were then electrophoresed at room temperature on a 4% polyacrylamide gel (19:1 acrylamide/bis) in 0.25× TBE buffer. Gels were dried and bands were visualized by autoradiography. Probes used for assays were Dm-70 bp 5'-CGATTGTAATTTTATATCGCCAGCAATGGACTGAAACGGTCCGCAACCTCTTCTTTACAACCTGGGTGAC-3' and the Pax5-binding site from the promoter of sea urchin H2a-2.2 5'-GGGTTGTGACGCAGCGGTGGGTGACGACTCCAGAGTCGACA-3'.

## DNase I FOOTPRINTING

*TaqI*–*SacII* fragment (130 bp) from the Dm segment, encompassing the detected Pax5-binding site, was labeled with <sup>32</sup>P-dCTP at *TaqI* end by a fill-in reaction with Klenow fragment to a specific activity greater than 10<sup>4</sup> cpm/ng of DNA. Probes were incubated for 20 min at room temperature with 20 µg of nuclear extract in a 50-µl reaction mixture containing 10 mM Tris pH 7.8, 14% glycerol, 57 mM KCl, 4 mM DTT, and 0.2 µg poly(dI-dC). DNase I (0.5–1 U; Promega) diluted in 50 mM MgCl<sub>2</sub>, 10 mM CaCl<sub>2</sub> was added for 1 min. The reaction was terminated by addition of 150 µl of a stop solution containing 200 mM NaCl, 20 mM EDTA, 1% SDS, and 5 µg yeast tRNA. DNA was extracted with phenol–chloroform, ethanol precipitated, dissolved in loading buffer (deionized Formamide – 5 mM EDTA), denatured for 10 min at 85°C and separated on a 6% polyacrylamide sequencing gel containing 7 M urea. Sequencing reactions performed using the Maxam and Gilbert procedure were run parallel to each probe.

## BISULFITE SEQUENCING

DNA was converted by bisulfite treatment using the EpiTect Bisulfite kit (Qiagen) and amplified by PCR with GoTaq (Promega) using the following primers; BisDm F 5'-TTGATAGATAGTTTAA GGGGTTTTT-3', BisDm R 5'-ATCTATCACATCTCTATTCTCTTCAAATTA-3', BisJκ2 F 5'-TTTTTGGAGAATGAATGTTAGTGTAATAAT-3', BisJκ2 R 5'-TAAACAATTTTCCCTCCTTAACAC-3'; ionJκ2 F 5'-(ion torrent A adapter)-(index)-GAAATGTTTAAAGAAGTAGGGTAGTTTGT-3'; ionJκ2 R 5'-(ion torrent P1 adapter)-CCCTCCTTAACACCTAATCTAAAAATAA-3'; ionJκ4 F 5'-(ion torrent A adapter)-(index)-ACCAAAAATAACTCATTTAACCAAATAT-3', ionJκ4R 5'-(ion torrent P1 adapter)-TGATTTTATGTTAGATTTGTGGGAR-3'. Amplicons were visualized on a 1.5% agarose gel, excised, and purified with the Qiaquick gel extraction kit (Qiagen). Amplicons intended for standard Sanger sequencing were TA cloned using pGEM-T easy kit (Promega). PCR with universal T7 and SP6 primers was performed on transformed colonies and correctly inserted clonal amplicons were sequenced by Sanger sequencing (ABI-Prism-3700). Samples amplified with ion torrent fusion primers were sequenced on an Ion Torrent Personal Genome machine (Invitrogen).

## CHROMATIN IMMUNOPRECIPITATION

IL-7-dependent pre-B cell cultures were made from the bone marrow of Igk<sup>WT/ΔDm</sup> mice as has been previously described (35). Cells were crosslinked with formaldehyde, chromatin extracted, and immunoprecipitated with an antibody directed against Pax5 (5 μg per 30 μg DNA) (SantaCruz). Semi-quantitative PCR was carried out on input DNA compared to immunoprecipitated DNA using primers specific for the Dm element and primers spanning the Dm deletion in order to test the enrichment on the WT and ΔDm alleles separately. PCR amplicons were visualized on an 8% polyacrylamide gel. Primers used: ΔDmChIP-F 5'-CCAAGAGATTGGATCGGAGA-3', ΔDmChIP-R 5'-CCATGACTTTTGCTGGCTGT-3'; WTDmChIP-F 5'-GGCCACGGTTTGTGAAGACA-3', WTDmChIP-R 5'-CAGGGTGAA CGCCAAATG-3', CD19-F 5'-GATTTGGAAGAGTGCCTACA-3', CD19-R 5'-GCCTGCCTCCTACTAAGGTA-3', β-actin-F 5'-CGCCATGGATGACGATATCG-3', β-actin-R 5'-CGAAGCCGGCTT TGCACATG-3'.

## SOMATIC HYPERMUTATION ANALYSIS OF PEYER'S PATCHES B CELLS

Peyer's patches (PP) were dissected from the small intestines of 4–6-month-old Igk<sup>WT/WT</sup>, Igk<sup>ΔDm/ΔDm</sup>, or Igk<sup>WT/ΔDm</sup> mice. PP from three to four mice were pooled for each experiment. PP were mashed through a 70 μm nylon mesh and washed with PBS to produce single cell suspensions. Cells were washed with PBS-0.5% BSA and labeled with PNA-FITC (Vector Labs) and αB220-PE (BD Bioscience). Germinal center B220<sup>+</sup>/PNA<sup>high</sup> B cells were sorted (FACSStar BD) to greater than 90% purity. WT and ΔDm rearranged Igk alleles were amplified with Vκ-Degenerate 5'-GTCCCTGCCAGGTTYAGTGGCAGTGGRTCWRRGGAC-3' and R3-1 5'-CAGACCCTGGTCTAATGGTTTGTAAACCACATGGG-3' primers using high fidelity PCR kit (Roche) with an initial denaturation of 4 min at 94°C, followed by 35 cycles of denaturation at 94°C for 15 s and annealing combined with elongation at 68°C for 2 min. 3' A-overhang nucleotides were added by 20 min incubation with Taq polymerase and ATP at 72°C. PCR fragments corresponding to Vκ-Jκ5 rearrangement of the WT and ΔDM (2.2 and 1.3 kb, respectively) were visualized on a 0.8% agarose gel, excised and purified with the QIAquick gel extraction kit and cloned into the TOPO-2.1 TA cloning vector (Invitrogen). Plasmids from single colonies were prepared and sequenced by Sanger sequencing (ABI-Prism-3700). Sequences were aligned to the Igk locus and mutations in the 188 bp region downstream of the Vκ-Jκ5 joint were analyzed.

## PYROSEQUENCING

RNA was extracted using tri-reagent (Sigma-Aldrich) from CD19<sup>+</sup> MACS sorted (Miltenyi Biotec) splenic cells of BALB/c/Cast Igk<sup>WT/WT</sup> and BALB/c/Cast Igk<sup>ΔDm/WT</sup> littermates, as well as control BALB/c and Cast mice. cDNA was prepared with mMLV reverse transcriptase (Promega) using random hexamer primers (Thermo Scientific). Rearranged Igk transcripts were amplified with Vκ-degenerate: 5'-GTCCCTGCCAGGTTYAGTGGCAGTGGRTCWRRGGAC-3' and biotinylated CκR-5'-GGGAA GCCTCCAAGACCTTA-3'. Resulting amplicons were visualized on a 1.5% agarose gel, excised and purified with the QIAquick gel extraction kit (Qiagen). Allelic distribution of

BALB/c/Cast transcripts was assessed by pyrosequencing on a PyroMark Q24 instrument (Qiagen) using Cκ-pyro primer 5'-ACATCAACTTCACCCAT-3'.

## LUCIFERASE REPORTER ASSAY

M12 cells were transiently transfected using the DEAE dextran method (36) with a luciferase reporter plasmid containing the minimal β-globin promoter TATA box (pTATA), without any additional regulatory elements or with insertions of the Dm element, iEk, or four NF-κB binding sites immediately upstream of the promoter. The cells were co-transfected with pβ-GAL to normalize for transfection efficiency. Luciferase activity was measured using the Luciferase Assay System (Promega) according to the manufacturer's instructions.

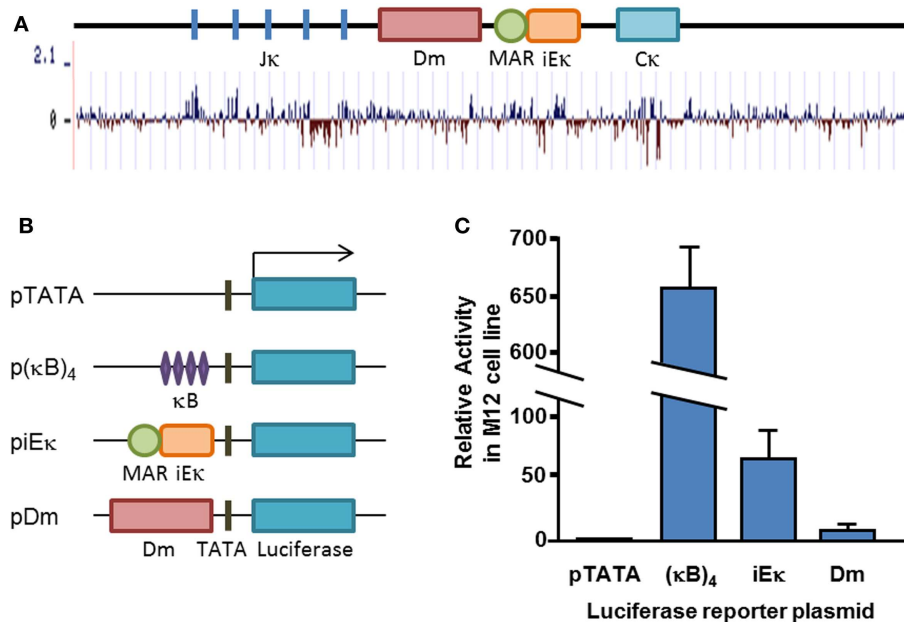
## RESULTS

### CHARACTERIZATION OF THE Dm ELEMENT

We have previously identified an element lying ~700 bp upstream of the iEk which facilitates demethylation of the Igk locus in cell culture, in cooperation with iEk (32). The element, designated Dm, is not part of the previously defined core iEk (**Figure 1A**). The Dm element, as determined by our previous experiments, spans ~1 kb and contains numerous areas which are conserved throughout different species (**Figure 1A**). The element itself contains a stretch of ~200 bp with the highest density of CpG sites found within the Igk locus. In order to see whether this element was transcriptionally active, we tested its functionality in an enhancer reporter assay. We compared its activity in a reporter plasmid to the well-characterized iEk (**Figure 1B**). Luciferase assays show that the Dm element acts only as a weak transcriptional enhancer which is about sevenfold weaker than the core intronic enhancer in M12 B cell lymphoma cells (**Figure 1C**), suggesting that the Dm element on its own does not exert its effect by direct transcriptional activation.

### Pax5 BINDING AT THE Dm ELEMENT

*Cis* regulatory elements, such as enhancers and promoters, convey their influence on cellular phenotypes by binding *trans* regulatory transcription factors, which mediate transcription and changes in chromatin structure. As the Igk locus is selectively active in B cells, starting from the pre-B cell stage, we speculated that the Dm element may bind B cell-specific transcription factors, thus mediating the changes it induces. Upon searching for potential binding sites for key B cell transcription regulators, we identified an area within the CpG-rich segment with remarkable similarity to the Pax5 consensus sequence (38) (**Figure 2A**). A 70-bp probe containing this sequence is shifted to a specific height when incubated with nuclear extracts from B lineage cells which have passed the pro-B cell stage, but not in other cell types tested in an electro-mobility shift assay (EMSA) (**Figure 2B**). These results clearly show that the binding of this protein is specific for the stages when the Igk locus is active. Notably, this specific shift can be attained using a fibroblast extract, which normally does not produce such a shift, by forced expression of Pax5 (**Figure 2C**), and titrated away by competition with a probe containing the Pax5-binding site of the H2a-2.2 promoter, strongly implying that indeed the Pax5 protein is binding at this site. When the nuclear extract is incubated



**FIGURE 1 | Conservation and transcriptional activity of the Dm element.** (A) Schematic map of the Ig $\kappa$  locus, drawn to scale. Basewise conservation across mammalian genomes is shown graphically, adapted from UCSC genome browser "Conservation" track. (B) Schematic map of transfected plasmid constructs. (C) Average relative luciferase activity in M12 cells

transfected with the indicated plasmids. Plasmids were co-transfected with a constitutive  $\beta$ -Gal-expressing plasmid and luciferase activity was normalized for transfection efficiency to  $\beta$ -Gal activity. Transfection and luciferase assay was carried out at least three times for each construct. Error bars represent the standard deviation of the luciferase activity.

with an antibody raised against the DNA-binding domain of Pax5 (designated A1), the shift on the EMSA gel disappears, whereas incubation with an antibody recognizing the Pax5 transactivation domain (designated A2) introduces a supershift, confirming that the 70-bp probe indeed specifically binds the Pax5 transcription factor (Figure 2D). DNase I footprinting using nuclei of the Pax5-expressing M12 B cell lymphoma cell line shows a definitive protection at the putative Pax5-binding site in comparison to S194 plasmacytoma cells which do not express Pax5 (Figure 2E). Interestingly, this specific footprint correlates precisely with the predicted Pax5-binding site. ChIP was performed on pre-B cells from Ig $\kappa$ <sup>WT/ $\Delta$ Dm</sup> mice (introduction of the  $\Delta$ Dm allele into mice is described in Section Characterization of Methylation, Rearrangement and B Cell Development in Dm Knockout Mice) with an antibody recognizing the Pax5 protein. While the Dm positive allele showed significant enrichment for Pax5, the deleted allele was not enriched for Pax5-binding (Figure 2F). These results indicate that Pax5 indeed binds in this region *in vivo* and that the binding is directly dependent on the presence of the Dm element. Altogether, the above described data shows that Pax5 specifically binds to the Dm element *in vivo*.

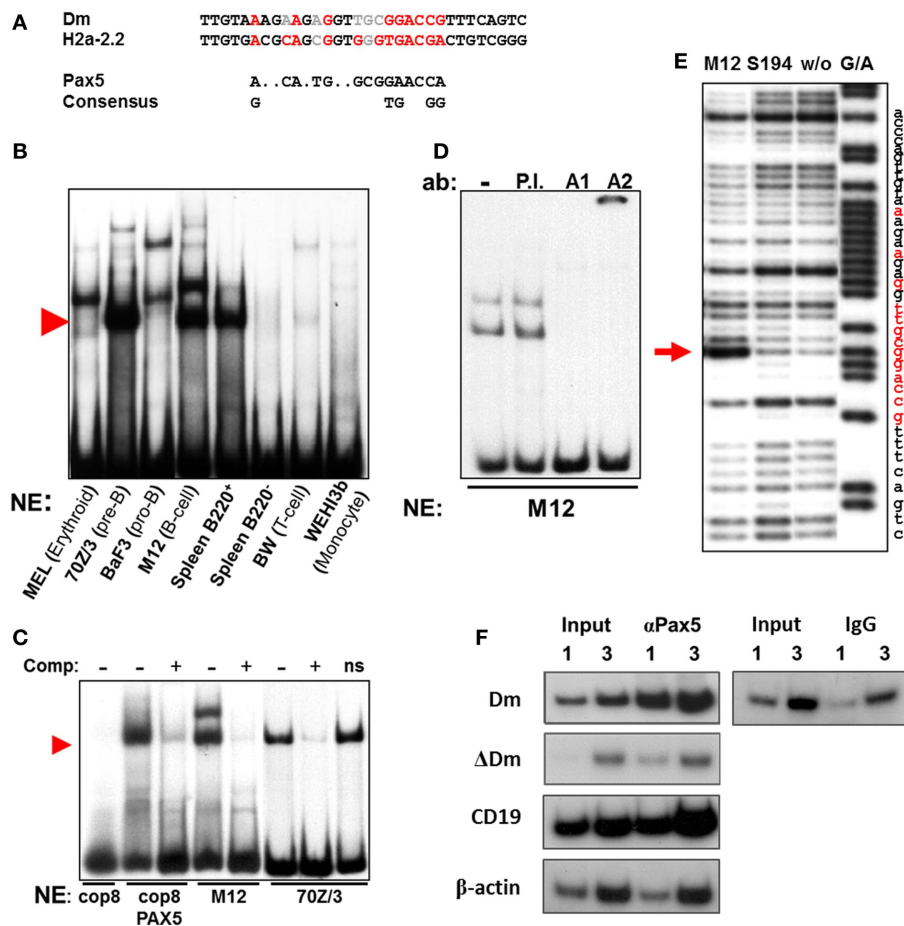
#### Dm FACILITATES DNA DEMETHYLATION OF Ig $\kappa$ TRANSGENES

We wished to further investigate the role of the Dm element in demethylation of the Ig $\kappa$  locus. In order to do so, we introduced a well-characterized transgene (34, 39) containing a pre-rearranged Ig $\kappa$  allele to mice, termed L $\kappa$  (Figure 3A). Two additional transgenic mice were produced with modified constructs, one containing a deletion of the entire Dm locus, termed L $\kappa$  $\Delta$ Dm, and the

second containing a deletion of the 70 bp region containing the Pax5-binding site, termed L $\kappa$  $\Delta$ 70 (Figure 3A). DNA from splenic B220<sup>+</sup> cells was assayed for the methylation of these transgenes by restriction analysis, which allows for simple differentiation between the transgenic and endogenous regions. Digestion with *KpnI* gave rise to a 4.3-kb fragment in the L $\kappa$  and L $\kappa$  $\Delta$ 70 transgenes and a 3.4-kb fragment in the L $\kappa$  $\Delta$ Dm transgene, whereas the endogenous locus yields a 15-kb fragment. These fragments were further digested with methylation-sensitive restriction enzymes *AciI* and *HhaI* (*HhaI* was not used to assess the L $\kappa$  $\Delta$ Dm state since the *HhaI* site is deleted in this transgene). The digested DNA was hybridized with a probe recognizing the MAR and iE $\kappa$  sequences. To assess the level of methylation, the amount of the undigested DNA was measured using a PhosphorImager. Interestingly, while the L $\kappa$  transgene was almost completely unmethylated, with only 8% of the DNA remaining undigested (Figures 3B,C), the L $\kappa$  $\Delta$ Dm transgene was highly methylated (73%) (Figures 3B,D), indicating that indeed the Dm element facilitates the hypomethylation of the Ig $\kappa$  locus in B cells. Notably, deletion of only 70 bp from the Dm in the L $\kappa$  $\Delta$ 70 transgene reduced the ability of the transgene to become unmethylated (50%) (Figures 3C,D).

Bisulfite analysis of the CpG-rich region surrounding the Pax5-binding site in the endogenous locus, L $\kappa$  and L $\kappa$  $\Delta$ 70 transgenes showed a picture that agrees quite nicely with the above results (the L $\kappa$  $\Delta$ Dm was not assayed in this manner, since this region is deleted within the transgene). These results take into consideration the difference between the methylation levels measured by bisulfite sequencing, which probes all CpG sites in the region, and the restriction analysis which measures the methylation only at





**FIGURE 2 | Pax5-binding within the Dm element. (A)** Comparison of the putative Pax5-binding site within the Dm element, known Pax5-binding site in sea urchin H2a-2.2 promoter and Pax5 consensus sequence. Bases which match the consensus sequence are marked in red, while bases which do not match are marked in gray. **(B)** Electro-mobility shift assay (EMSA) of *TaqI*–*Bst*EII 70 bp fragment from the Dm element, containing the putative Pax5-binding site, with the indicated hematopoietic cell type nuclear extracts. Red arrowhead indicates B cell lineage-specific shift. **(C)** Competition EMSA. Unlabeled H2a-2.2 probe competes with radioactive *TaqI*–*Bst*EII Dm probe and reduces shift in B lineage cell extracts (M12, 70Z/3), and in extracts of fibroblast cells (cop8) transfected with a Pax5 expression vector. Nonsense unlabeled probe (ns) is unable to compete with Dm probe. Red arrowhead indicates specific shift. **(D)** Supershift assay with *TaqI*–*Bst*EII Dm probe, which was incubated with antibodies (ab) raised against the DNA-binding site of Pax5 (A1), the transactivation domain (A2),

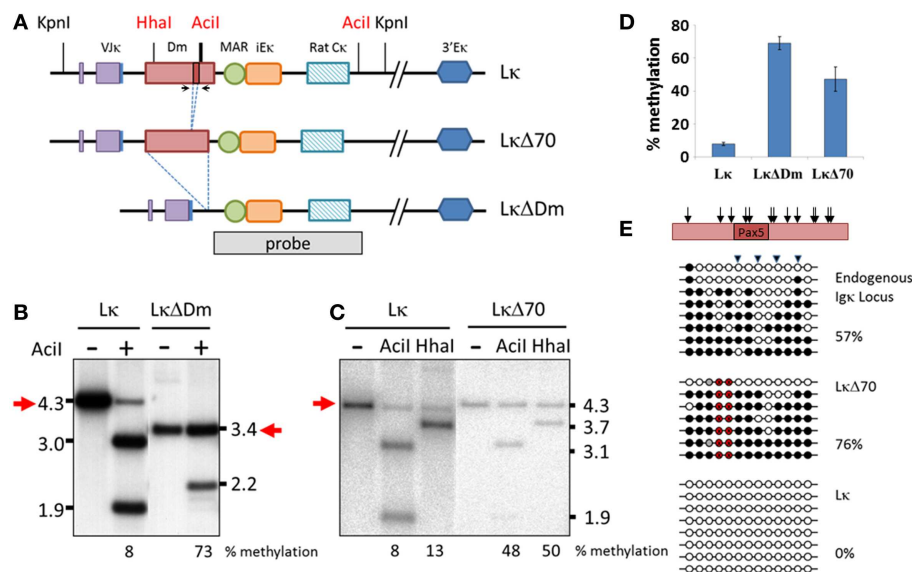
or general rabbit antisera (P.I.). **(E)** DNase I footprinting assay on end-labeled *TaqI*–*Sac*II probe from the Dm element. Labeled probe was incubated with nuclear extracts from the indicated cell types. A control Maxam and Gilbert reaction (G/A) was run in parallel. Footprint specific to Pax5-expressing M12 cells is indicated with a red arrow. Location of the putative Pax5-binding site is marked in red on the nucleotide sequence. **(F)** Chromatin immunoprecipitation (ChIP) enrichment of Pax5 at the Dm element in Igκ<sup>wt/ΔDm</sup> cultured pre-B cells. Enrichment was measured by semi-quantitative PCR via comparison of the input DNA (Input) to the immunoprecipitated fraction (αPax5), using primers specific to the Dm element (Dm) and the deleted allele (ΔDm). One and three times the amount of PCR template were run in parallel to ensure linearity. Positive (CD19 promoter) and negative (β-actin promoter) controls for Pax5-binding were analyzed in parallel to ensure specificity of the enrichment. ChIP with a non-specific antibody (IgG) does not enrich the Dm element.

the sites which correspond to the digestion site. The endogenous locus is close to 50% methylated, as expected from a region which undergoes monoallelic demethylation (Figure 3E). The LkΔ70 transgene is 76% methylated, while the Lk transgene is completely unmethylated (Figure 3E). In order to see how these results correlate with the restriction analysis, the percent of sequences which would be protected from *Aci*I digestion was assessed. Fifty-seven percent of the LkΔ70 sequences remain protected, supporting the restriction analysis results. These experiments clearly show that the Dm element contributes to the demethylation of the Igκ locus

*in vivo*, results that support previously published data obtained from cell culture systems.

#### CHARACTERIZATION OF METHYLATION, REARRANGEMENT, AND B CELL DEVELOPMENT IN Dm KNOCKOUT MICE

Given the results in transgenic mice, we generated a knockout mouse in which the entire Dm element in the endogenous locus was replaced with a LoxP-flanked Neo gene which was then excised from the genome (Figure 4A; Figure S1 in Supplementary Material). We assessed the methylation pattern of the Igκ locus by



**FIGURE 3 | Effect of Dm element on methylation of Ig $\kappa$  locus in transgenic mice. (A)** Schematic map of the transgenes and probe used in experiments. Bold *Acil* line indicates multiple *Acil* restriction sites within close proximity. **(B,C)** Southern blot assay assessing methylation state of L $\kappa$ , L $\kappa$  $\Delta$ Dm **(B)** and L $\kappa$  $\Delta$ 70 **(C)** transgenes in B220<sup>+</sup> splenic B cells. DNA was digested with *KpnI* either alone or with methylation sensitive enzymes *Acil* or *HhaI* and hybridized with the indicated probe. Methylation levels were quantified by phosphorImager. **(D)** Graphical representation of the average methylation levels at *Acil* sites of the indicated transgenes, as quantified by phosphorImager. Methylation was

measured in DNA from two to four independent founder mice for each transgene. Error bars represent standard deviation of the methylation levels. **(E)** Bisulfite analysis of CpGs flanking the Pax5-binding sequence of the Dm element. Triangles mark CpGs recognized by *Acil* restriction enzyme. Black circles signify methylated, white signify unmethylated, gray signify undetermined, and red signify CpGs deleted in transgene. Percentage of methylated CpGs is noted. Relative location of the CpGs within the Dm element is indicated with arrows on schematic representation of the genomic region. Location of primers is marked by arrows in **(A)**.

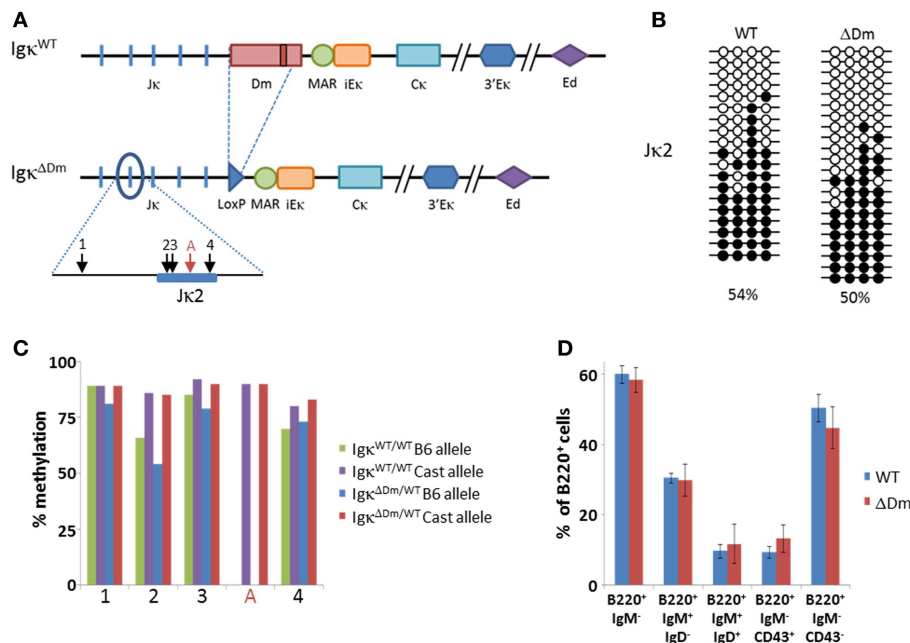
bisulfite analysis of the J $\kappa$ 2 fragment in *ex vivo* mature B cells. Surprisingly, given the strong phenotype in transgenic mice, no significant difference was seen between the methylation levels of Ig $\kappa$ <sup>WT/WT</sup> and Ig $\kappa$  <sup>$\Delta$ Dm/ $\Delta$ Dm</sup> mice (**Figure 4B**).

We then proceeded to investigate whether the methylation patterns at the Ig $\kappa$  locus are affected by deletion of the Dm element in the pre-B cell stage, which is the very first stage in which demethylation of the locus is detected. To this end, Ig $\kappa$  <sup>$\Delta$ Dm</sup> mice were bred onto a Rag1<sup>-/-</sup> background, effectively blocking rearrangement of the Ig $\kappa$  locus and differentiation to the mature B cell stage. Expression of a pre-rearranged IgH transgene was ensured in order to allow the cells to express the pre-BCR and differentiate to the pre-B cell stage. These mice were further bred with Rag1<sup>-/-</sup> *M. castaneus* mice, which contain an intact Dm element, thus allowing distinction between the WT and the Dm-deleted alleles based on the strain-specific DNA polymorphisms. *Ex vivo* CD19<sup>+</sup> bone marrow cells were purified from B6/Cast Rag1<sup>-/-</sup> IgH<sup>+</sup> Ig $\kappa$  <sup>$\Delta$ Dm/WT</sup> and B6/Cast Rag1<sup>-/-</sup> IgH<sup>+</sup> Ig $\kappa$ <sup>WT/WT</sup> mice and the methylation of the J $\kappa$ 2 and J $\kappa$ 4 segments was determined by high-throughput sequencing. We did not, however, detect significant differences in levels of methylation between the WT and  $\Delta$ Dm alleles (**Figure 4C**, Figure S2 in Supplementary Material). Taken together, we find that, while the Dm element plays a role in demethylation of the Ig $\kappa$  locus in transgenes, this role is not translated to the endogenous locus, probably due to redundancy of the many enhancers of the locus, not all of which are present in the transgene.

We explored the possibility that the Dm element may affect other developmental processes pertaining the Ig $\kappa$  locus and normal B cell development, as has been observed for *cis* regulatory elements in the locus such as the enhancers. There was no significant difference seen in the levels of rearrangement of the WT versus  $\Delta$ Dm allele, as assessed by FACS and pyrosequencing analyses (Figure S3 in Supplementary Material). The pyrosequencing results also indicate that the level of Ig $\kappa$  transcription is not changed by the deletion of the Dm element, supporting the above described results showing that the Dm element is a weak transcriptional enhancer. The B cell development in the bone marrow of Ig $\kappa$  <sup>$\Delta$ Dm/ $\Delta$ Dm</sup> mice appeared normal, with proportions of pro-, pre-, immature, and mature B cells similar to those of WT mice (**Figure 4D**, Figure S4 in Supplementary Material). Overall, these results indicate that, in the endogenous locus, deletion of the Dm element does not curtail these early stages of B cell development.

#### EFFECT OF Dm ELEMENT ON SHM

We investigated whether deletion of the Dm element affects a later stage of Ig $\kappa$  maturation, specifically the process of SHM in activated B cells. Levels of SHM in Ig $\kappa$ <sup>WT/WT</sup> mice versus Ig $\kappa$  <sup>$\Delta$ Dm/ $\Delta$ Dm</sup> mice were examined, and a significant drop in amount of mutations in the germinal center B220<sup>+</sup> PNA<sup>high</sup> B cells from Peyer's patches of the Dm negative mice was observed (**Figures 5A,B**). In order to rule out mouse to mouse variation, which could potentially give rise to such an effect, SHM in heterozygous Ig $\kappa$ <sup>WT/ $\Delta$ Dm</sup>



**FIGURE 4 | Effect of deletion of the Dm element at the endogenous locus on Ig $\kappa$  methylation and B cell development in the bone marrow.**

(A) Schematic map of the endogenous Ig $\kappa$  locus in wild-type (WT) and Dm knockout ( $\Delta$ Dm) mice. Relative locations of CpGs in Jk2 region are indicated with arrows. CpG present only in Castaneous (Cast) strain is marked with a red arrow. (B) Bisulfite analysis of CpGs at the Jk2 region in splenic CD19<sup>+</sup> B cells from WT and  $\Delta$ Dm mice. Black circles signify methylated CpGs, white circles signify unmethylated CpGs. Percentage of methylated CpGs is noted. (C) Bisulfite analysis by high-throughput sequencing of Jk2 region from

CD19<sup>+</sup> bone marrow pre-B cells from Rag1<sup>-/-</sup> C57BL/6/Castaneous IgH-3H9-Tg mice with or without a deletion of the Dm element on the C57BL/6 (B6) allele. Copies (1600–3000) of each CpG from each genotype were analyzed. Alleles were differentiated by strain-specific polymorphic sites within the amplified regions. The methylation state of each CpG is summarized graphically. (D) Summary of proportions of B cell populations within bone marrows of WT and  $\Delta$ Dm mice. Error bars mark standard deviation. Six mice were analyzed in each group. Representative FACS plots can be seen in Figure S4 in Supplementary Material.

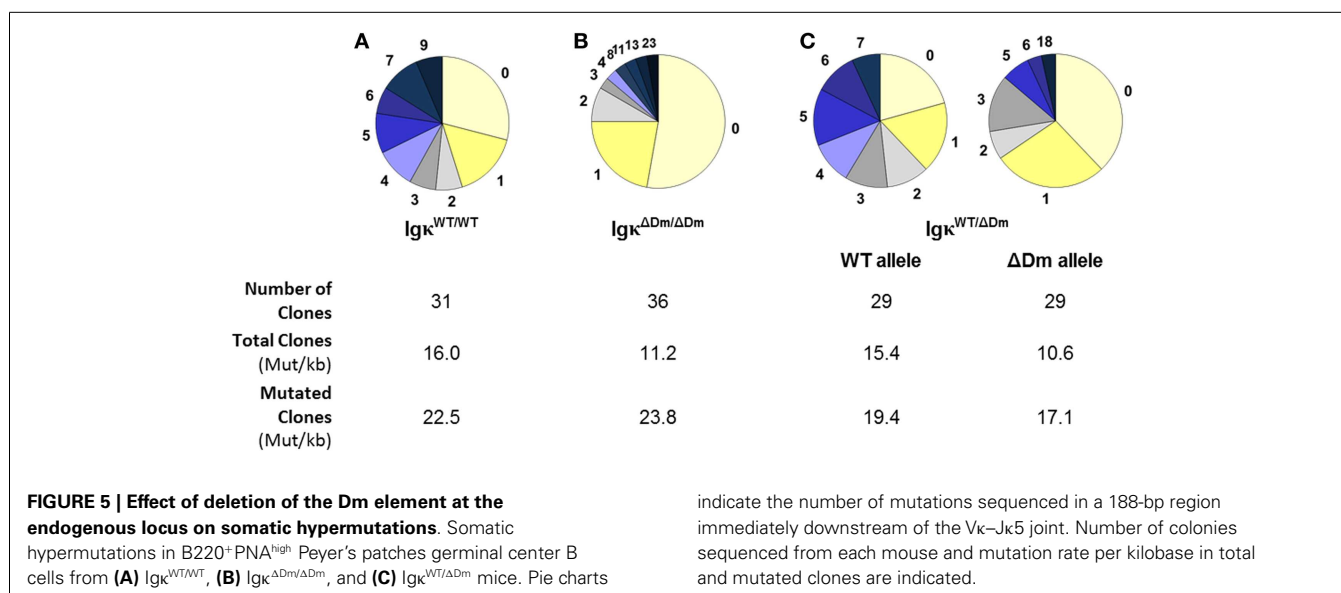
mice was assessed. Here too, the proportion of mutations on the  $\Delta$ Dm allele was lower than on the WT allele (Figure 5C). As a control, a similar number of colonies were sequenced from B220<sup>+</sup> PNA<sup>low</sup> cells, with no mutations detected (data not shown). While the average number of mutations is lower in the  $\Delta$ Dm allele, sequences which have undergone SHM do so at an efficiency similar to the WT allele, as seen when examining the mutation rate in total sequences versus rate in mutated sequences (Figure 5), suggesting that the Dm element affects the recruitment but not the processivity of the machinery involved in SHM. These results indicate that the Dm element, which is immediately adjacent to the intronic MAR and iE $\kappa$ , helps promote SHM. This is particularly notable, as deletion of the MAR/iE $\kappa$  region on its own has no discernable effect on the normal SHM process (18). Our results clearly show that the Dm element contributes to proper SHM at the Ig $\kappa$  locus, a role which has not been previously attributed to the intronic enhancer region.

## DISCUSSION

In this paper, we characterized a novel *cis* regulatory element situated within the J $\kappa$ –C $\kappa$  intron of the Ig $\kappa$  locus. This sequence was originally identified as an element which lies adjacent to iE $\kappa$  and contributes to its demethylating activity, as deletion of either element was sufficient to abolish demethylation in a cell culture system (32). In our present study, we find that the Dm element is

necessary for hypomethylation of the Ig $\kappa$  locus of the L $\kappa$  transgene *in vivo*, but is dispensable for the demethylation of the endogenous locus. The apparent discrepancy between the phenotype in these two cases may be due to the fact that the transgene contains the sequences in the Ig $\kappa$  locus up to the 3'E $\kappa$ , but does not include the Ed. The three Ig $\kappa$  enhancers work cooperatively and, to a certain extent, redundantly to activate and demethylate the locus. Previous studies have shown that deletion of any single enhancer has only a small effect on the developmentally regulated DNA demethylation, whereas the combined lack of two enhancers abolishes the demethylation process (16, 19). Another difference between the transgene and the endogenous locus is that the transgene contains a pre-rearranged Ig $\kappa$ . It is possible that the Dm element only affects the demethylation when the locus is in a rearranged configuration, but not in the germline conformation. In this study, we see that the deletion of the Dm sequence, which is not part of the core iE $\kappa$ , greatly impedes the demethylation process in the transgene, indicating that the Dm element contributes to the activity of iE $\kappa$ , possibly as a co-enhancer. As the Dm is only a weak transcriptional enhancer as a solitary element, it is the combined activity with neighboring *cis* acting elements which gives rise to the full activity.

The mechanisms by which genomic loci undergo targeted demethylation have long been shrouded in mystery (40). Findings from recent years have pointed to the Tet family of enzymes as



possible catalysts of the demethylation process, *via* oxidation of the methyl group into a hydroxymethyl moiety (41). When acting as a demethylation intermediate, the hydroxymethylated cytosine is then either passively diluted during DNA replication (42), as it is not recognized by the methylation maintenance machinery (43, 44), or, conversely, is actively excised from the genome and replaced with an unmethylated nucleotide (45, 46). Targeting of Tet proteins to specific genomic loci is sufficient to induce local demethylation (47, 48). Tet2 has been implicated in the active demethylation of tissue-specific genes in postmitotic human monocytes (49). Additionally, Tet2 has been found to bind PU.1 (50) and EBF1 (51) in the hematopoietic system. A recent report has uncovered a different strategy to induce demethylation by which DNMT1, the maintenance DNA methyltransferase, is sequestered from specific genomic loci by binding non-coding RNA. This prevents the placement of methyl groups on DNA during replication and, in turn, brings about passive demethylation of a defined region (52). It is still unclear what mechanism is implemented by the *cis* regulatory elements to demethylate the Igκ locus during B cell development, especially since deletion of Tet2, the strongest Tet candidate in the immune system, causes leukemia in mice (53–55), which masks many of the tissue-specific effects that may occur as a result. As methylation is a strong barrier to the rearrangement process (8), future studies can address this issue.

We have identified a sequence within the Dm element which binds the B cell lineage specifier Pax5. This site is bound by Pax5 starting with the pre-B cell stage, up to mature B cells, but is unbound in Baf3 pro-B cells, where the Igκ locus is not yet activated and made accessible for rearrangement, nor in plasma cells where Pax5 expression is down-regulated. It should be noted, though, that the Baf3 pro-B cell line tested here does not express Pax5 (56), whereas most pro-B cells do, and as such we are unable to rule out the possibility that Pax5 is already bound at the pro-B cell stage. This is the first report, to our knowledge, of a Pax5-binding site within the Igκ locus which binds Pax5 at the time of locus activation. Previous reports have located sites in 3'Ek (24) and

in the K-I-K-II (29, 30) regulatory elements in which Pax5 plays a repressive role and where binding is lost upon Igκ locus activation. The new site we report is particularly interesting, considering that Pax5 is known to be directly necessary for Igκ locus activation and κ<sub>0</sub> germline transcription in pre-B cells (31). We find that Pax5-binding in the vicinity of the Jκ-Cκ intron is dependent on the presence of the Dm element and that the Pax5-binding site contributes to the demethylating capabilities of the Dm element. While this clearly cannot be the only Pax5-binding site, since ΔDm pre-B cells maintain their full ability to rearrange the Igκ locus, this site highlights the potency of this B cell identity protein in one more area of B cell development.

It should be noted that the sequence of the Pax5-binding site within the Dm element is conserved among rodents, but not in the human-Igκ locus, though other aspects, such as the CpG-dense region, are. This is not the only aspect which differs between the human and murine counterparts of the Igκ locus. For example, the Sis element, a transcriptional silencer which has been shown to recruit the Igκ locus to the pericentromeric heterochromatin in mice, is not conserved in the human locus (57). It stands to reason that regulation of the human and murine Igκ loci may differ somewhat, as the strongly biased usage of the κ versus λ chain seen in mice (where 95% of mature B cells express the κ chain) is not present in humans, which have a ratio of 60:40 of κ versus λ usage (58). This could be due to differences in the RR of the human and murine Igκ locus that may contribute to this phenomenon.

While the deletion of the Dm element did not, on its own, affect the methylation status of the endogenous Igκ locus, nor the relative amount of the deleted allele which underwent rearrangement, we observed a decrease in the levels of SHM on Igκ alleles lacking the Dm element. The role of the Dm element in facilitating SHM appears to be independent of the iEk/MAR region, since the combined deletion of the iEk and MAR elements has no perceptible effect on SHM (18). The lower level of SHM does not appear to be the result of lower levels of Igκ transcription, since deletion of the Dm element does not lower the levels of Igκ RNA



observed in mature B cells. Deletion of the Dm element appears to cause inefficient recruitment of the mutating machinery, but once the machinery is in place, the mutation efficiency is similar to the WT locus. The element may therefore function by efficiently recruiting the mutation machinery to the locus, possibly by key regulators such as Pax5 which are bound to the Dm element. Pax5 itself has a known role in SHM by activating the transcription of the *Aicda* gene, encoding the AID protein, which is the deaminase responsible for SHM (59, 60). It may be that Pax5 plays more than one role in SHM induction. The role of the Dm element in SHM fits in well with its location, which is almost immediately adjacent to the V $\kappa$ –J $\kappa$  rearranged region which is the hotspot for SHM.

In conclusion, we have characterized the Dm sequence as an element that regulates the Igk locus during different stages of B cell development. The Dm is both a team player, cooperating with the three characterized enhancers to demethylate the locus for rearrangement, as well as an element that affects the locus in its own right in allowing efficient SHM. This report adds to our understanding of the complex regulation of the Igk locus, which undergoes many drastic changes during development and must be fine-tuned for each developmental stage.

## AUTHOR CONTRIBUTIONS

Rena Levin-Klein, Andrei Kirillov, and Chaggai Rosenbluh designed the experiments, did the research, and interpreted the results. Howard Cedar and Yehudit Bergman directed the study. Rena Levin-Klein and Yehudit Bergman wrote the manuscript.

## ACKNOWLEDGMENTS

We would like to thank Prof. Meinrad Busslinger for providing antibodies, Pax5 expression vector, and his scientific expertise regarding Pax5, Prof. Michael Neuberger for providing the original Lk transgenic mouse and Lk construct, Prof. Klaus Rajewsky and Dr. Raul Mostoslavsky for assistance in generating Igk<sup>ΔDm</sup> mice, Gidon Toperoff for assistance with Pyrosequencing, and Adam Spiro for assistance with bioinformatics analysis. This work was supported by research grants from the Israel Academy of Sciences (Howard Cedar, Yehudit Bergman), NIH (Yehudit Bergman), the Israel Cancer Research Foundation (Howard Cedar, Yehudit Bergman), and the USA-Israel Binational Science Foundation (Yehudit Bergman).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at <http://www.frontiersin.org/journal/10.3389/fimmu.2014.00240/abstract>

## REFERENCES

- Gellert M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem* (2002) **71**:101–32. doi:10.1146/annurev.biochem.71.090501.150203
- Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell* (1985) **40**(2):271–81. doi:10.1016/0092-8674(85)90141-2
- Constantinescu A, Schlissel MS. Changes in locus-specific V(D)J recombinase activity induced by immunoglobulin gene products during B cell development. *J Exp Med* (1997) **185**(4):609–20. doi:10.1084/jem.185.4.609
- Stanhope-Baker P, Hudson KM, Shaffer AL, Constantinescu A, Schlissel MS. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. *Cell* (1996) **85**(6):887–97. doi:10.1016/S0092-8674(00)81272-6
- Mostoslavsky R, Singh N, Kirillov A, Pelanda R, Cedar H, Chess A, et al. Kappa chain monoallelic demethylation and the establishment of allelic exclusion. *Genes Dev* (1998) **12**(12):1801–11. doi:10.1101/gad.12.12.1801
- Nakase H, Takahama Y, Akamatsu Y. Effect of CpG methylation on RAG1/RAG2 reactivity: implications of direct and indirect mechanisms for controlling V(D)J cleavage. *EMBO Rep* (2003) **4**(8):774–80. doi:10.1038/sj.embor.embor904
- Goldmit M, Ji Y, Skok J, Roldan E, Jung S, Cedar H, et al. Epigenetic ontogeny of the Igk locus during B cell development. *Nat Immunol* (2005) **6**(2):198–203. doi:10.1038/ni1154
- Goldmit M, Schlissel M, Cedar H, Bergman Y. Differential accessibility at the kappa chain locus plays a role in allelic exclusion. *EMBO J* (2002) **21**(19):5255–61. doi:10.1093/emboj/cdf518
- Odegard VH, Schatz DG. Targeting of somatic hypermutation. *Nat Rev Immunol* (2006) **6**(8):573–83. doi:10.1038/nri1896
- Fraenkel S, Mostoslavsky R, Novobrantseva TI, Pelanda R, Chaudhuri J, Esposito G, et al. Allelic 'choice' governs somatic hypermutation in vivo at the immunoglobulin kappa-chain locus. *Nat Immunol* (2007) **8**(7):715–22. doi:10.1038/ni1476
- Queen C, Baltimore D. Immunoglobulin gene transcription is activated by downstream sequence elements. *Cell* (1983) **33**(3):741–8. doi:10.1016/0092-8674(83)90016-8
- Meyer KB, Neuberger MS. The immunoglobulin kappa locus contains a second, stronger B-cell-specific enhancer which is located downstream of the constant region. *EMBO J* (1989) **8**(7):1959–64.
- Liu ZM, George-Raizen JB, Li S, Meyers KC, Chang MY, Garrard WT. Chromatin structural analyses of the mouse Igkappa gene locus reveal new hypersensitive sites specifying a transcriptional silencer and enhancer. *J Biol Chem* (2002) **277**(36):32640–9. doi:10.1074/jbc.M204065200
- Xu Y, Davidson L, Alt FW, Baltimore D. Deletion of the Ig kappa light chain intronic enhancer/matrix attachment region impairs but does not abolish V kappa J kappa rearrangement. *Immunity* (1996) **4**(4):377–85. doi:10.1016/S1074-7613(00)80251-4
- Gorman JR, van der Stoep N, Monroe R, Cogne M, Davidson L, Alt FW. The Ig(kappa) enhancer influences the ratio of Ig(kappa) versus Ig(lambda) B lymphocytes. *Immunity* (1996) **5**(3):241–52. doi:10.1016/S1074-7613(00)80319-2
- Inlay M, Alt FW, Baltimore D, Xu Y. Essential roles of the kappa light chain intronic enhancer and 3' enhancer in kappa rearrangement and demethylation. *Nat Immunol* (2002) **3**(5):463–8. doi:10.1038/ni790
- Xiang Y, Garrard WT. The Downstream Transcriptional Enhancer, Ed, positively regulates mouse Ig kappa gene expression and somatic hypermutation. *J Immunol* (2008) **180**(10):6725–32. doi:10.4049/jimmunol.180.10.6725
- Inlay MA, Gao HH, Odegard VH, Lin T, Schatz DG, Xu Y. Roles of the Ig kappa light chain intronic and 3' enhancers in Igk somatic hypermutation. *J Immunol* (2006) **177**(2):1146–51. doi:10.4049/jimmunol.177.2.1146
- Zhou X, Xiang Y, Garrard WT. The Igkappa gene enhancers, E3' and Ed, are essential for triggering transcription. *J Immunol* (2010) **185**(12):7544–52. doi:10.4049/jimmunol.1002665
- Inlay MA, Lin T, Gao HH, Xu Y. Critical roles of the immunoglobulin intronic enhancers in maintaining the sequential rearrangement of IgH and Igk loci. *J Exp Med* (2006) **203**(7):1721–32. doi:10.1084/jem.20052310
- Zhou X, Xiang Y, Ding X, Garrard WT. A new hypersensitive site, HS10, and the enhancers, E3' and Ed, differentially regulate Igkappa gene expression. *J Immunol* (2012) **188**(6):2722–32. doi:10.4049/jimmunol.1102758
- Cockerill PN, Garrard WT. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs next to the enhancer in a region containing topoisomerase II sites. *Cell* (1986) **44**(2):273–82. doi:10.1016/0092-8674(86)90761-0
- McDevitt DC, Perkins L, Atchison ML, Nikolajczyk BS. The Ig kappa 3' enhancer is activated by gradients of chromatin accessibility and protein association. *J Immunol* (2005) **174**(5):2834–42. doi:10.4049/jimmunol.174.5.2834
- Shaffer AL, Peng A, Schlissel MS. In vivo occupancy of the kappa light chain enhancers in primary pro- and pre-B cells: a model for kappa locus activation. *Immunity* (1997) **6**(2):131–43. doi:10.1016/S1074-7613(00)80420-3
- Roque MC, Smith PA, Blasquez VC. A developmentally modulated chromatin structure at the mouse immunoglobulin kappa 3' enhancer. *Mol Cell Biol* (1996) **16**(6):3138–55.



26. Lazorchak AS, Schlissel MS, Zhuang Y. E2A and IRF-4/Pip promote chromatin modification and transcription of the immunoglobulin kappa locus in pre-B cells. *Mol Cell Biol* (2006) **26**(3):810–21. doi:10.1128/MCB.26.3.810-821.2006
27. Inlay MA, Tian H, Lin T, Xu Y. Important roles for E protein binding sites within the immunoglobulin kappa chain intronic enhancer in activating V $\kappa$  rearrangement. *J Exp Med* (2004) **200**(9):1205–11. doi:10.1084/jem.20041135
28. Maitra S, Atchison M. BSAP can repress enhancer activity by targeting PU.1 function. *Mol Cell Biol* (2000) **20**(6):1911–22. doi:10.1128/MCB.20.6.1911-1922.2000
29. Sato H, Wang D, Kudo A. Dissociation of Pax-5 from KI and KII sites during kappa-chain gene rearrangement correlates with its association with the underphosphorylated form of retinoblastoma. *J Immunol* (2001) **166**(11):6704–10. doi:10.4049/jimmunol.166.11.6704
30. Tian J, Okabe T, Miyazaki T, Takeshita S, Kudo A. Pax-5 is identical to EBB-1/KLP and binds to the VpreB and lambda5 promoters as well as the KI and KII sites upstream of the Jkappa genes. *Eur J Immunol* (1997) **27**(3):750–5. doi:10.1002/eji.1830270325
31. Sato H, Saito-Ohara F, Inazawa J, Kudo A. Pax-5 is essential for kappa sterile transcription during Ig kappa chain gene rearrangement. *J Immunol* (2004) **172**(8):4858–65. doi:10.4049/jimmunol.172.8.4858
32. Lichtenstein M, Keini G, Cedar H, Bergman Y. B cell-specific demethylation: a novel role for the intronic kappa chain enhancer sequence. *Cell* (1994) **76**(5):913–23. doi:10.1016/0092-8674(94)90365-4
33. Casellas R, Shih TA, Kleinsteinfeld M, Rakonjac J, Nemazee D, Rajewsky K, et al. Contribution of receptor editing to the antibody repertoire. *Science* (2001) **291**(5508):1541–4. doi:10.1126/science.1056600
34. Meyer KB, Sharpe MJ, Surani MA, Neuberger MS. The importance of the 3'-enhancer region in immunoglobulin kappa gene expression. *Nucleic Acids Res* (1990) **18**(19):5609–15. doi:10.1093/nar/18.19.5609
35. Farago M, Rosenbluh C, Tevlin M, Fraenkel S, Schlesinger S, Masika H, et al. Clonal allelic predetermination of immunoglobulin-kappa rearrangement. *Nature* (2012) **490**(7421):561–5. doi:10.1038/nature11496
36. Gulick T. Transfection using DEAE-dextran. *Curr Protoc Cell Biol* (2003) **Chapter 20**:Unit20.4. doi:10.1002/0471143030.cb2004s19
37. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol* (1975) **98**(3):503–17. doi:10.1016/S0022-2836(75)80083-0
38. Czerny T, Schaffner G, Busslinger M. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev* (1993) **7**(10):2048–61. doi:10.1101/gad.7.10.2048
39. Sharpe MJ, Milstein C, Jarvis JM, Neuberger MS. Somatic hypermutation of immunoglobulin kappa may depend on sequences 3' of C kappa and occurs on passenger transgenes. *EMBO J* (1991) **10**(8):2139–45.
40. Ooi SK, Bestor TH. The colorful history of active DNA demethylation. *Cell* (2008) **133**(7):1145–8. doi:10.1016/j.cell.2008.06.009
41. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* (2009) **324**(5929):930–5. doi:10.1126/science.1170116
42. Inoue A, Zhang Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* (2011) **334**(6053):194. doi:10.1126/science.1212483
43. Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res* (2012) **40**(11):4841–9. doi:10.1093/nar/gks155
44. Otani J, Kimura H, Sharif J, Endo TA, Mishima Y, Kawakami T, et al. Cell cycle-dependent turnover of 5-hydroxymethyl cytosine in mouse embryonic stem cells. *PLoS One* (2013) **8**(12):e82961. doi:10.1371/journal.pone.0082961
45. Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* (2011) **145**(3):423–34. doi:10.1016/j.cell.2011.03.022
46. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* (2011) **333**(6047):1303–7. doi:10.1126/science.1210944
47. Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, et al. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat Biotechnol* (2013) **31**(12):1137–42. doi:10.1038/nbt.2726
48. Chen H, Kazemier HG, de Groote ML, Ruiters MH, Xu GL, Rots MG. Induced DNA demethylation by targeting ten-eleven translocation 2 to the human ICAM-1 promoter. *Nucleic Acids Res* (2014) **42**(3):1563–74. doi:10.1093/nar/gkt1019
49. Klug H, Schmidhofer S, Gebhard C, Andreessen R, Rehli M. 5-Hydroxymethylcytosine is an essential intermediate of active DNA demethylation processes in primary human monocytes. *Genome Biol* (2013) **14**(5):R46. doi:10.1186/gb-2013-14-5-r46
50. de la Rica L, Rodriguez-Ubrea J, Garcia M, Islam AB, Urquiza JM, Hernando H, et al. PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome Biol* (2013) **14**(9):R99. doi:10.1186/gb-2013-14-9-r99
51. Guilhamon P, Eskandarpour M, Halai D, Wilson GA, Feber A, Teschendorff AE, et al. Meta-analysis of IDH-mutant cancers identifies EBF1 as an interaction partner for TET2. *Nat Commun* (2013) **4**:2166. doi:10.1038/ncomms3166
52. Di Ruscio A, Ebralidze AK, Benoukrat T, Amabile G, Goff LA, Terragni J, et al. DNMT1-interacting RNAs block gene-specific DNA methylation. *Nature* (2013) **503**(7476):371–6. doi:10.1038/nature12598
53. Ko M, Bandukwala HS, An J, Lamperti ED, Thompson EC, Hastie R, et al. Ten-eleven-translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc Natl Acad Sci U S A* (2011) **108**(35):14566–71. doi:10.1073/pnas.1112317108
54. Li Z, Cai X, Cai CL, Wang J, Zhang W, Petersen BE, et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood* (2011) **118**(17):4509–18. doi:10.1182/blood-2010-12-325241
55. Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* (2011) **20**(1):11–24. doi:10.1016/j.ccr.2011.06.001
56. Sigvardsson M, Clark DR, Fitzsimmons D, Doyle M, Akerblad P, Breslin T, et al. Early B-cell factor, E2A, and Pax-5 cooperate to activate the early B cell-specific mb-1 promoter. *Mol Cell Biol* (2002) **22**(24):8539–51. doi:10.1128/MCB.22.24.8539-8551.2002
57. Liu Z, Widlak P, Zou Y, Xiao F, Oh M, Li S, et al. A recombination silencer that specifies heterochromatin positioning and ikaros association in the immunoglobulin kappa locus. *Immunity* (2006) **24**(4):405–15. doi:10.1016/j.immuni.2006.02.001
58. Popov AV, Zou X, Xian J, Nicholson IC, Bruggemann M. A human immunoglobulin lambda locus is similarly well expressed in mice and humans. *J Exp Med* (1999) **189**(10):1611–20. doi:10.1084/jem.189.10.1611
59. Gonda H, Sugai M, Nambu Y, Katakai T, Agata Y, Mori KJ, et al. The balance between Pax5 and Id2 activities is the key to AID gene expression. *J Exp Med* (2003) **198**(9):1427–37. doi:10.1084/jem.20030802
60. Tran TH, Nakata M, Suzuki K, Begum NA, Shinkura R, Fagaras S, et al. B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. *Nat Immunol* (2010) **11**(2):148–54. doi:10.1038/ni.1829

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 January 2014; accepted: 08 May 2014; published online: 23 May 2014.

Citation: Levin-Klein R, Kirillov A, Rosenbluh C, Cedar H and Bergman Y (2014) A novel Pax5-binding regulatory element in the Ig $\kappa$  locus. *Front. Immunol.* 5:240. doi: 10.3389/fimmu.2014.00240

This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Levin-Klein, Kirillov, Rosenbluh, Cedar and Bergman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The interplay between chromatin and transcription factor networks during B cell development: who pulls the trigger first?

Mohamed Amin Choukrallah<sup>1\*</sup> and Patrick Matthias<sup>1,2\*</sup>

<sup>1</sup> Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

<sup>2</sup> Faculty of Sciences, University of Basel, Basel, Switzerland

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

Subbarao Bondada, University of Kentucky, USA

Kay L. Medina, Mayo Clinic, USA  
Marcus R. Clark, The University of Chicago, USA

## \*Correspondence:

Mohamed Amin Choukrallah and  
Patrick Matthias, Friedrich Miescher  
Institute for Biomedical Research,  
Maulbeerstrasse 66, Basel CH-4058,  
Switzerland  
e-mail: mohamed-amin.choukrallah@fmi.ch; patrick.matthias@fmi.ch

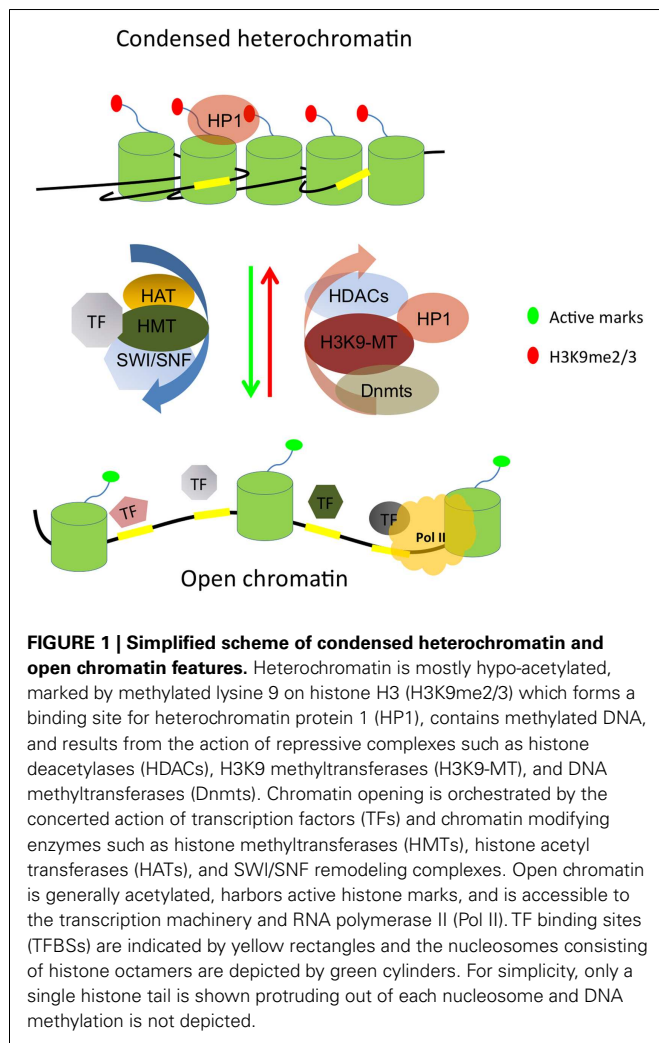
All mature blood cells derive from hematopoietic stem cells through gradual restriction of their cell fate potential and acquisition of specialized functions. Lineage specification and cell commitment require the establishment of specific transcriptional programs involving the activation of lineage-specific genes and the repression of lineage-inappropriate genes. This process requires the concerted action of transcription factors (TFs) and epigenetic modifying enzymes. Within the hematopoietic system, B lymphopoiesis is one of the most-studied differentiation programs. Loss of function studies allowed the identification of many TFs and epigenetic modifiers required for B cell development. The usage of systematic analytical techniques such as transcriptome determination, genome-wide mapping of TF binding and epigenetic modifications, and mass spectrometry analyses, allowed to gain a systemic description of the intricate networks that guide B cell development. However, the precise mechanisms governing the interaction between TFs and chromatin are still unclear. Generally, chromatin structure can be remodeled by some TFs but in turn can also regulate (i.e., prevent or promote) the binding of other TFs. This conundrum leads to the crucial questions of who is on first, when, and how. We review here the current knowledge about TF networks and epigenetic regulation during hematopoiesis, with an emphasis on B cell development, and discuss in particular the current models about the interplay between chromatin and TFs.

**Keywords: hematopoiesis, B cell development, transcription factors, chromatin regulators, pioneer transcription factors**

## INTRODUCTION

The character of a cell type is defined by its specific transcriptional program, which is regulated by transcription factors (TFs) that bind DNA cis-regulatory elements (cis-REs) to activate or repress defined set of genes. Cis-REs refer to loci that regulate the expression of genes located in the same molecule of DNA. They are composed of binding sites for TFs that recognize a specific nucleotide sequence and therefore act as trans-acting factors. Promoters and enhancers are the two major types of cis-REs in eukaryotes. At the DNA sequence level, the repertoire of cis-REs is identical in all cell types. Therefore, the transcriptional programs specific to each cell lineage must be the consequence of the repertoire of TFs expressed in a given cell that select genes for transcriptional activation or repression. However, the same TFs can be equally expressed in different cell types but have distinct binding profiles, indicating that the interaction between the TFs and their cognate sequences is not sufficient to explain the action of TFs and their transcriptional output. Indeed, in addition to the DNA sequence recognition, TFs occupancy strongly depends on chromatin structure and epigenetic modifications which provide an additional layer of gene regulation and establish heritable cellular memories.

Chromatin consists of repeating units of nucleosomes, comprising histone octamers (containing two copies each of H2A, H2B, H3, and H4) around which 147 bp of DNA are wrapped (1). Multiple residues within the tails and the globular domains of histones can undergo post-translational modifications (PTMs), including acetylation, methylation, phosphorylation, ubiquitination, and sumoylation. These PTMs are catalyzed by a variety of histone-modifying enzymes that have been classified in two major groups, the writers such as histone acetyl transferases (HATs) and histone methyltransferases (HMTs) and the erasers such as histone deacetylases (HDACs) and histone demethylases (KDMs) (2) (Figure 1). Histone modifications act combinatorially to regulate transcriptional activity; some histone modifications are associated with transcriptional activation while others are associated with distinct mechanisms of transcriptional repression. For example, tri-methylation on lysine 4 of histone H3 (H3K4me3) is mainly associated with active promoters; in contrast, the mono methylation of the same residue (H3K4me1) is a hallmark of poised and active enhancers, while H3K27ac marks exclusively active enhancers and promoters. The best studied repressive histone modifications are the methylation of lysines 9 and 27 of histone H3 (H3K9 and H3K27), which are respectively associated with



heterochromatin and polycomb-group (PcG) proteins-mediated repression.

DNA methylation provides yet an additional mechanism for gene regulation. It is an efficient repressive DNA modification that occurs at the fifth position of cytosine (5-mC), mostly in the context of CpG dinucleotides (3) and is associated with transcriptional repression through two general mechanisms. First, DNA methylation can directly inhibit the binding of proteins important for transcription initiation, such as TFs and others. Moreover, methylated DNA can recruit proteins containing a methylated DBD, which may interfere with transcription by co-recruitment of repressors such as HDACs [reviewed in Ref. (4)]. Most of the genome is depleted of CpGs except for CpG islands, which represent ca. 60% of mammalian promoters and are largely unmethylated (5). DNA methylation is catalyzed by three enzymes: the maintenance DNA methyltransferase Dnmt1, which ensures that already methylated residues are faithfully maintained during DNA replication (6), and the *de novo* methyltransferases Dnmt3A and Dnmt3B which can add methyl groups to non-methylated CpG residues (7). DNA methylation is dynamic and also reversible: removal of methyl groups can occur through active or passive

mechanisms. The latter is due to the absence of methylation by Dnmt1 of newly synthesized DNA during replication. In contrast, active DNA demethylation corresponds to the reaction that leads to the removal of the methyl group from 5-mC residues independent of DNA replication. Active DNA demethylation has been a controversial subject as many mechanisms were proposed to explain this process and the putative demethylases could not be identified in a conclusive manner [reviewed in Ref. (8)]. However, it is now well accepted that the dioxygenases Tet1 and Tet2 catalyze DNA demethylation through the conversion of 5-mC to hydroxymethyl cytosine (5-hmC) (9, 10).

Additional mechanisms involved in epigenetic regulation are contributed by chromatin remodeling complexes (CRC) and diverse kinds of non-coding RNAs. Chromatin remodelers are ATP-dependent complexes that regulate DNA accessibility by modifying nucleosome positioning and conformation. They can be divided into four groups: the SWI/SNF, ISWI, CHD, and INO80 families of remodelers [reviewed in Ref. (11, 12)]. In addition, long or short non-coding RNAs can influence chromatin and gene expression, for example by mediating inactivation of one chromosome (X inactivation by Xist RNA), opening up loci or helping to define boundaries of chromatin domains [reviewed by Mercer et al. (13)].

These different mechanisms of histone modifications, DNA methylation, chromatin remodeling, and non-coding RNAs play a central role in shaping chromatin structure, which in turn affects the interaction between TFs and their cognate binding sites. Conversely, the binding of TFs triggers a chain of events, often leading to changes in local chromatin properties. Indeed, TFs can interact with and recruit many chromatin modifying or remodeling complexes to their target loci. Thus, establishing chromatin structure requires TF activity and TF activity depends on chromatin structure. This reciprocal interplay raises a major question: how is the communication between TFs and chromatin regulated and which additional cellular signals feed into this complex network during development and cellular differentiation?

Understanding the mutual and interdependent interactions between TFs and chromatin features and their impact on gene regulation in a developmental system requires a biological paradigm where successive differentiation stages can easily be identified and isolated. In this regard, hematopoiesis provides a powerful system to study epigenetic and transcriptional dynamics. B cells derive from hematopoietic stem cells (HSCs) through a multistep differentiation program. HSCs have both self-renewal and multipotency capacities. The precise balance of these properties is essential to maintain the HSC pool size throughout animal life. HSCs initially give rise to multipotent progenitors (MPPs) that lose self-renewal capacity but keep the ability to generate early progenitors of lymphoid and myeloid lineages. Lymphoid lineage consists of B, T, and natural killer (NK) cells while myeloid lineage contains macrophages (M), granulocytes (G), erythrocytes (E), and megakaryocytes (Mk). The exact branching point between lymphoid and myeloid lineages as well as the differentiation potential of progenitor populations is still matter of some debate [reviewed in Ref. (14)]. The identification of common lymphoid progenitors (CLPs) (15) and common myeloid progenitors (CMPs) (16) supports the model that lymphoid and myeloid lineages follow

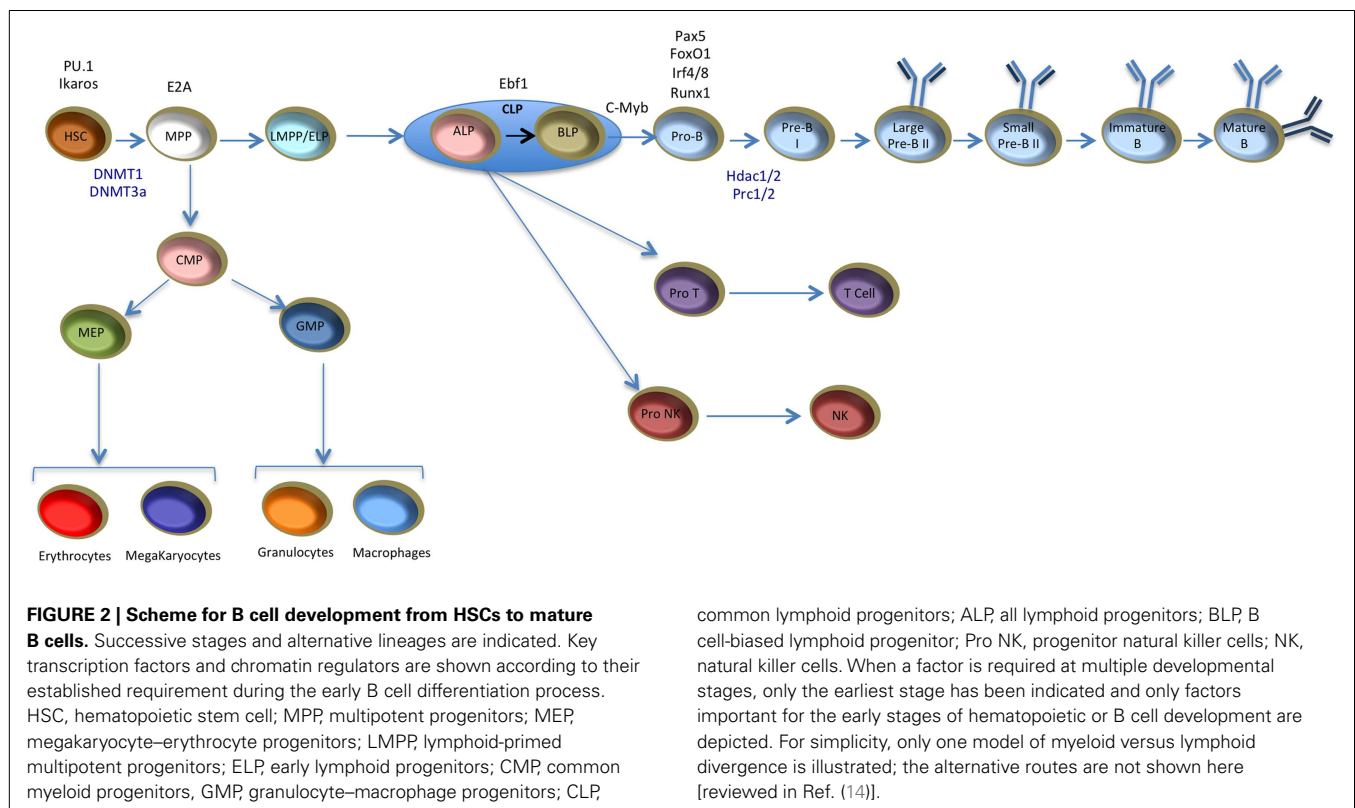
distinct developmental paths from MPPs. This model was challenged by the identification of the lymphoid-primed multipotent progenitors (LMPPs) that lose M<sub>KE</sub> potential but keep lymphoid and GM potential (17, 18). Another study showed that the MPP compartment contains a subpopulation of cells with strong lymphoid potential and weak myeloid colony-forming activity (19). These cells, called early lymphoid progenitors (ELPs) start to express recombination-activating gene 1 (Rag1) and Rag2 and initiate the immunoglobulin heavy chain (IgH) rearrangement (19). ELPs are thought to precede the CLP stage. Recently, it was shown that the CLP compartment contains two distinct subpopulations: all lymphoid progenitors (ALPs) and B cell-biased lymphoid progenitors (BLPs) (20). ALPs retain full lymphoid potential, whereas BLPs behave essentially as B cell progenitors (20). Mature B cells derive from BLPs through sequential differentiation steps that can be defined by five major stages that are phenotypically and functionally distinct: pro-B, pre-BI, large and small pre-B II, and immature B cells (21) (**Figure 2**). Early B cell differentiation is intimately connected to the DNA rearrangement of Ig genes, the so-called V(D)J recombination, in order to generate functional Ig molecules. Pro-B cells, first express the pan-B cell marker B220 and this coincides with entry into the B cell lineage. Next, pre-BI cells express the CD19 gene and complete recombination of the IgH diversity (D<sub>H</sub>) to joining (J<sub>H</sub>) segments and the next stage sees the generation of IgH V(D)J alleles [reviewed in Ref. (21)]. This allows expression of the rearranged heavy chain which assembles with the surrogate light chain to form the pre-B cell receptor (pre-BCR), a crucial checkpoint in B cell development (22). If cells pass this functional test they can go on to the next developmental stage,

small pre-BII cells, where the Ig light chain rearranges and allows for the formation and exposure at the cell surface of a functional Ig molecule, the BCR. Finally, immature cells can then leave the bone marrow (BM) and enter the periphery (22).

The generation of immature and mature B cells from early precursors is a progressive process, every step of which is characterized by a specific transcription program involving the activation, repression, or maintenance of distinct sets of gene expression patterns. This genetic regulation results from the concerted action of ubiquitous and lineage-specific TFs as well as epigenetic modifiers. Proper and timely recombination of the Ig loci is essential for normal progression through B cell development and is highly dependent on chromatin structure, DNA methylation, and also expression of various RNAs across the Ig locus. In particular, the accessibility model, first posited by Frederick Alt and colleagues, highlighted the importance of “sterile” transcripts which originate from unrearranged Ig gene segments and make their chromatin accessible to the recombination enzymes RAG1 and RAG2 [reviewed in Ref. (23)]. Thus, B cell development presents an extraordinarily complex and dynamic system to study the establishment and maintenance of transcriptional and epigenetic networks.

### KEY TRANSCRIPTION FACTORS ESSENTIAL FOR B CELL DEVELOPMENT

Loss of function studies using mouse models have identified many TFs important for distinct stages during B cell development and a particular emphasis has been put on early B cell specification and commitment. Prominent among those are E2A, Ebf1, and Pax5, as well as other TFs acting downstream and upstream to these



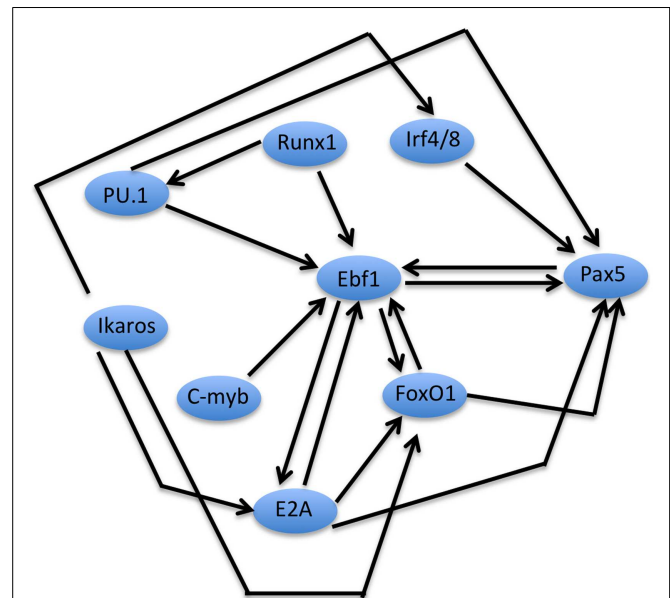


factors. Some of these TFs such as Ebf1 and Pax5 are restricted to the B cell lineage while others such as Ikaros, PU.1, E2A, and FoxO1 are also involved in other lineage fate determination.

The expression of these TFs is temporally regulated; e.g., Ikaros, PU.1, and E2A are expressed in the very early progenitors including HSCs and MPPs before the commitment to the lymphoid branch, Ebf1 and FoxO1 are expressed at the CLP stage under the control of E2A (24) and Pax5 expression is induced by the concerted action of Ebf1, FoxO1, and E2A in committed pro-B cells. The sequential expression and activity of these TFs suggests a hierarchy in their action. Yet, the transcriptional regulation of early B cell development is not a simple hierarchical cascade, as many of these TFs act in a cooperative manner and directly regulate the expression of other TFs, involving both positive and negative feedback loops leading to a complex cross-regulatory network (25) (**Figure 3**).

### PU.1

If one considers a hierarchical classification of the TFs involved in B cell development, PU.1 and Ikaros come on the top of the regulatory pyramid. PU.1 (encoded by *Spi-1*, *Sfpi-1*) belongs to the ETS family of TFs, its expression was thought to be restricted to the hematopoietic lineage, but was recently also detected in adipocytes (26). Within the hematopoietic system, PU.1 activity is essential for the development of lymphoid cells as well as macrophages and neutrophils (27). Disruption of the PU.1 DBD in mouse prevents the commitment of MPPs toward the lymphoid lineage (27). PU.1 is expressed in HSCs (28), lymphoid, and myeloid progenitors (16) as well as in fully mature and functional cells (29). This broad expression pattern indicates that PU.1 is not only required for cell differentiation but also plays a role in the function of the specialized hematopoietic cells. The expression of PU.1 in many hematopoietic lineages raised the question about its mechanism of action and the rules that determine the interaction between this TF and its binding sites in different cellular and physiological contexts. Genome-wide mapping of PU.1 binding sites in macrophages and B cells revealed that PU.1 is enriched at transcription start sites (TSSs), but the majority of binding sites were found at inter- and intra-genic sites (30) indicating a role of PU.1 in regulating both transcription initiation and enhancer function. Interestingly, PU.1 binding at TSSs exhibits a high correlation between macrophages and B cells; in contrast, binding sites at distal regulatory elements are highly cell type-specific. Motif analysis of cell type-specific PU.1 binding sites revealed that PU.1 binds in vicinity of lineage-specific TFs: B cell-specific PU.1 binding sites are enriched in E2A, Ebf1, OCT, and NF- $\kappa$ B motifs, while macrophage-specific sites are enriched in C/EBP and AP-1 motifs (30). These findings strongly suggest that the cell type-specific function of PU.1 is partly due to its collaborative interaction with other lineage-specific TFs. The role of PU.1 in shaping the enhancer repertoire in hematopoietic cells will be further discussed in a later section of this review. Interestingly, PU.1 action was shown to depend critically on its expression level and involves a tight dose-dependent control. PU.1 shows low to medium expression level in LT-HSCs and exhibits varied levels in progenitors and mature cells; e.g., PU.1 is weakly expressed in erythroid and T cells and shows intermediate levels in B cells, in contrast, it is highly expressed in macrophages and neutrophils (31). Importantly, this graded expression has a critical



**FIGURE 3 | An early B cell specification module.** Schematized network of inter-dependent TFs regulating early B cell development. The scheme depicts TFs that have been implicated in the control of early B cell specification and/or commitment (for clarity, factors such as Spi-B, OBF1, or NF- $\kappa$ B, which are important at later stages of B cell development, are not depicted here). Direct positive regulation between two factors at the transcriptional level is indicated by the corresponding arrows. Note that the scheme does not imply biochemical interactions (e.g., complex formation) between the factors, although they may also take place in some cases.

role in specifying the different lineages: by artificial expression of PU.1 in PU.1-deficient progenitors, it could be demonstrated that moderate PU.1 levels promote B cell development, while high PU.1 expression promotes macrophage differentiation and at the same time blocks B cell development (32).

### IKAROS

The zinc finger factor Ikaros (encoded by the *ikzf1* gene) also plays a critical role during early lymphoid lineage specification. Ikaros was proposed to promote the differentiation of pluripotent HSCs into the lymphocyte pathways: mutational disruption of the Ikaros DNA-binding domain (DBD) leads to an early block in lymphopoiesis before the commitment to the lymphoid restricted stages (33). However, another study showed that Ikaros is dispensable for the transition from HSCs to LMPPs, but is rather required for the progression of LMPPs into the lymphoid lineages (34). Recently, Ikaros was found to be required for the induction of lymphoid lineage priming in HSCs and for the repression of self-renewal and multipotency genes after HSC differentiation (35). Ikaros is also involved in later stages of B cell development, where it promotes heavy-chain gene rearrangements by inducing expression of the RAG1 and RAG2 genes, as well as by controlling accessibility of the variable gene segments and compaction of the IgH locus (36). Furthermore, Ikaros was recently shown to be required for the differentiation of large pre-B to small pre-B cells and for transcription and rearrangement of the IgL locus (37). Ikaros functions either as a transcriptional activator or repressor by



recruiting various CRC including SWI/SNF and Mi-2/nucleosome remodeling and deacetylase (Mi-2/NuRD) to DNA regulatory elements and to pericentromeric heterochromatin (38–42).

## E2A

E2A (encoded by *tcf3* with two splice variants, E12 and E47) is a helix–loop–helix TF essential for B cell differentiation (43, 44). E2A-null mutant mice fail to generate LMPPs and lack B cells (43). E2A acts synergistically with PU.1 and is required for Ebf1 and FoxO1 expression at the CLP stage (45, 46). Genome-wide mapping experiments in B cell progenitors (Ebf1<sup>-/-</sup> and Rag2<sup>-/-</sup>) showed that E2A binds both TSSs and putative enhancers (24) and is required to induce H3K4me1 deposition at enhancer elements in concert with PU.1 (30).

## EARLY B CELL FACTOR 1

Early B cell factor 1 (Ebf1) belongs to the EBF/COE family of TFs (47). EBF/COE family members contain an N-terminal DBD with an atypical zinc knuckle domain (H–X3–C–X2–C–X5–C), a TF immunoglobulin (TIG/IPT) domain, a helix–loop–helix–loop–helix (HLHLH) domain and a carboxy-terminal transactivation domain (48). The HLHLH domain was found to be important for the dimerization of EBF1 (48). EBF1 is essential for B cell specification (49) and commitment (50). Ebf1 acts in concert with E2A, FoxO1, and other TFs to regulate the expression of many genes required for B cell development including TFs such as FoxO1 and Pax5 (51); the latter in turn binds to *Ebf1* enhancers and increases its expression, thereby leading to a positive feedback loop between these two factors (24, 52, 53). Ebf1 can also act as a repressor; indeed, it was shown that Ebf1 prevents Id2- and Id3-mediated inhibition of the E47 isoform of E2A by downregulating the expression of their mRNA (54).

## Pax5

Pax5 acts downstream of Ebf1, its expression is under the control of a cohort of TFs including PU.1, Ebf1, FoxO1, IRF4, and IRF8 (55). Pax5 is essential for B cell commitment (56) and maintenance of B cell identity through activation of B cell-specific genes and repression of lineage-inappropriate genes (57). Deletion of Pax5 in mature B cells leads to the de-differentiation to lymphoid progenitors, which can differentiate into functional T cells (58). The role of Pax5 in regulating gene expression will be discussed in more detail in a later section.

## FoxO1

The forkhead TF FoxO1 plays an important role during B cell development. FoxO1 was found to be critical at several stages of B cell differentiation (59). Early deletion of FoxO1 causes a substantial block at the pro-B cell stage due to a failure to express the IL-7 receptor- $\alpha$  chain. FoxO1 inactivation in late pro-B cells results in an arrest at the pre-B cell stage due to impaired expression of Rag1 and Rag2 (59), which are direct targets of FoxO1 (60). In addition, deletion of FoxO1 in peripheral B cells leads to reduced number of lymph node B cells due to down regulation of L-selectin and defect in class-switch recombination (59).

## c-Myb AND Runx1

B cell development also depends on many other TFs such as for example c-Myb and Runx1. Deletion of c-Myb in mice leads

to a block at the pre–pro-B cell stage which is accompanied with impaired expression of the  $\alpha$ -chain of the IL-7 receptor and Ebf1 (61). Deletion of Runx1 also causes a developmental block at the pro-B cell stage accompanied by reduced expression of E2A, Ebf1, and Pax5. Furthermore, Runx1 directly binds the *Ebf1* promoter and this binding is critical for *Ebf1* activation; indeed, Runx1-deficient pro-B cells were shown to harbor excessive amounts of the repressive histone mark H3K27me3 in the *Ebf1* proximal promoter. Interestingly, retroviral transduction of Ebf1, but not Pax5, into Runx1-deficient progenitors restores B cell development (62). It was also shown that Runx1 controls the expression of PU.1 via direct interaction with its upstream regulatory element (URE) (63).

As discussed above, many of the TFs critical for early B cell development directly regulate each other's expression, positively or negatively, by binding to cis-REs in their corresponding genes. This inter-dependent network forms a B specification module, which has EBF1 at its center, which in concert with Ikaros, E2A, IRF4/8, and FoxO1, positively activates expression of Pax5, thus locking B cell development (Figure 3).

## EPIGENETIC REGULATORS INVOLVED IN HEMATOPOIESIS AND B CELL DEVELOPMENT

In addition to the TFs, many epigenetic regulators are crucial for hematopoiesis and/or B cell development. Among those, PcG proteins play an important role in this system. Mammalian cells contain two major PcG complexes, PRC1 and PRC2. PRC2 contains SUZ12, EED, and EZH1 or EZH2. EZH proteins are HMTs that catalyze the di- and tri-methylation of histone H3K27 (64). PRC1 contains RING1, CBX, PHC, and BMI1 or MEL18 [reviewed in Ref. (65)]. PRC1 recognizes and binds H3K27me3 via its subunit CBX, while RING1 mono-ubiquitylates histone H2A at lysine 119 (H2AK119ub1) (61, 62). The H2AK119ub1 mark is thought to play a role in inhibiting RNA polymerase II (pol II) elongation (66). The H3K27me3 mark is associated with the silencing of many key developmental regulatory genes, such as Hox homeotic genes and many others [reviewed in Ref. (67)].

Many PcG deficiencies correlate with defective development and/or activation of lymphocytes. For example, inactivation of Bmi1 or mel-18 causes a severe block in B cell development that leads to B cell lymphopenia (68, 69). By contrast, deficiency in Cbx2 does not affect lymphocyte development but alters splenic B cell response to lipopolysaccharide (LPS) (70). Conditional knockout studies targeting members of the polycomb machinery highlighted the critical role of these enzymatic complexes in the hematopoietic system. Bmi1 is the most studied PRC1 subunit in hematopoiesis. Depletion of Bmi1 leads to impaired self-renewal capacity of HSCs due to the de-repression of two major cell cycle regulators: Ink4a (p16) and Arf (p19) (71). Bmi1 directly binds and repress the promoters of these genes and the deletion of both *Ink4a* and *Arf* genes restores the self-renewal capacity of *Bmi1*<sup>-/-</sup> HSCs (72). Moreover, *Bmi1*<sup>-/-</sup> mice have a BM microenvironment that is severely defective in supporting hematopoiesis. In this case however, the deletion of both *Ink4a* and *Arf* genes did not significantly restore the impaired BM microenvironment (72). Bmi1 is also involved in the repression of Ebf1 and Pax5 in HSCs and MPPs. Depletion of Bmi1 causes aberrant expression of these two genes, leading to premature lymphoid lineage

specification (73). Another PRC1 subunit, Ring1b, was also found to be critical for adult hematopoiesis. Mice deficient for Ring1b in hematopoietic cells develop a hypocellular BM that unexpectedly contains an enlarged, hyperproliferating compartment of immature cells, with an intact differentiation potential. These defects are associated with differential upregulation of cyclin D2 and Ink4a (74). Controlled expression of PRC2 components is also important for hematopoiesis. Several studies have highlighted the role of Ezh1 and Ezh2 in embryonic and adult HSCs. Loss of Ezh2 severely impairs fetal HSC self-renewal without affecting the function of adult stem cells present in the BM (75). In addition, EZH2 was also found to have a crucial role in early B cell development and in rearrangement of the IgH gene (66).

Early B cell development also requires HDACs activity (76). Targeted deletion of the major class I HDACs, HDAC1 and 2 showed that B cell development requires the presence of at least one of these two enzymes. When both enzymes are deleted, B cell development is dramatically impaired at the large pre-BII stage with a strong cell cycle block in the G1 phase accompanied by the induction of apoptosis. In contrast, elimination of HDAC1 and HDAC2 in mature resting B cells is not deleterious; however, when these cells are induced to proliferate cell cycle block and apoptosis ensue. These data indicate that the role of HDAC1 and 2 during early B cell development is at least partially linked to cell cycle control (76). The potential role of HDACs in controlling other processes in B cells and other hematopoietic lineages remains to be elucidated.

The activity of DNA methyltransferases is also crucial for hematopoiesis. Conditional deletion of the maintenance DNA methyltransferase Dnmt1 in HSCs leads to impaired self-renewal capacity and prevents HSCs from giving rise to hematopoietic progenitors (77). Based on the initial studies, loss of the *de novo* DNA methyltransferases, Dnmt3a or Dnmt3b alone was thought to have no impact on HSC function (78); by contrast, loss of both together was reported to abolish self-renewal without affecting differentiation capacity (78). However, a more recent study reported that Dnmt3a-null HSCs exhibit upregulation of multipotency genes and downregulation of differentiation factors. The progeny of Dnmt3a-deficient HSCs exhibit global hypomethylation and impaired repression of HSC-specific genes. These data highlighted the important role of Dnmt3a in the repression of HSC genes in order to enable proper cell differentiation (79).

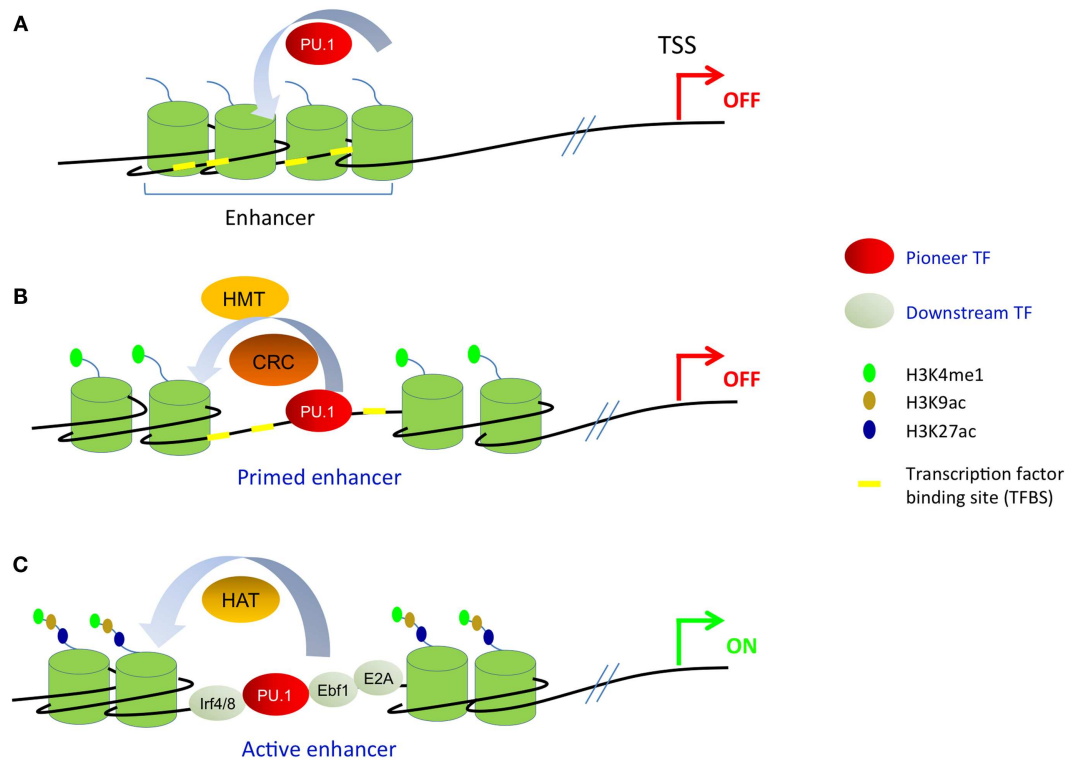
V(D)J recombination of immunoglobulin genes is thought to be regulated by changes in the accessibility of target sites, such as modulation of methylation. *In vitro* experiments showed that specific methylation within the heptamer of recombination signal sequences markedly reduces V(D)J cleavage without inhibiting RAG1/RAG2–DNA complex formation (80). Recent investigations of the IgH locus recombination showed that the diversity ( $D_H$ ) and joining ( $J_H$ ) gene segments are methylated prior to recombination, in contrast the  $DJ_H$  product is demethylated.  $DJ_H$  junctional demethylation is restricted to B cells and requires the  $E\mu$  enhancer, located within the intronic region of the IgH locus (81). However, it is unclear whether the demethylation is required for  $DJ_H$  junction or whether it is simply the consequence of the DNA recombination. Earlier experiments had shown that loss of methylation of the kappa light chain locus is not sufficient to

activate recombination in cultured pre B cells lacking Dnmt1 (82). Cd19-cre mediated deletion of the *de novo* DNA methyltransferase Dnmt3a and Dnmt3b, failed to identify a critical role for these enzymes in B cell development (83). Cd19-cre is expected to induce the deletion of targeted genes from pre B cells onward (84). Thus, this study, strongly suggest that Dnmt3a and b are dispensable to the progression from pre B cells to mature B cells. Overall, these studies suggest that the maintenance DNA methyltransferase Dnmt1 is required at all the stages of hematopoiesis, whereas the *de novo* DNA methyltransferases Dnmt3a and 3b are required only at the very early stages and become dispensable at later stages. However, additional studies will be needed to fully test these assumptions.

## INTERPLAY BETWEEN CHROMATIN LANDSCAPE AND TF ACTIVITY DURING B CELL DEVELOPMENT

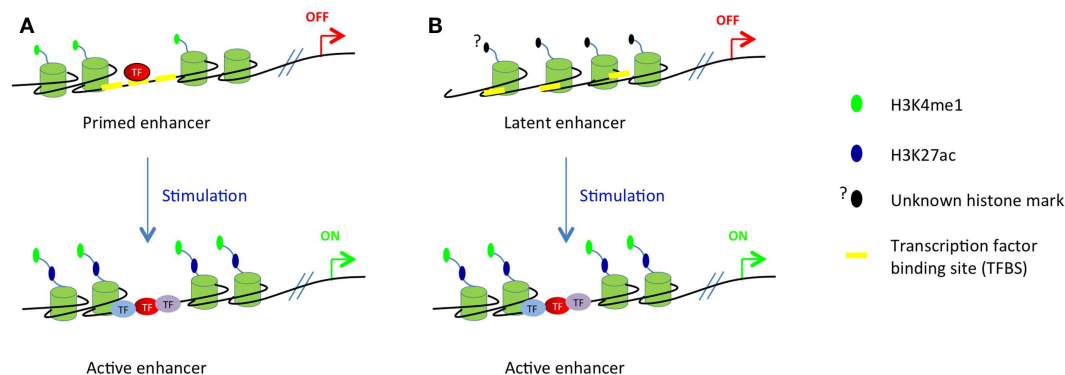
The progression of MPPs toward specialized cells is thought to be accompanied by extensive epigenetic reprogramming. In the recent years, genome-wide technologies have been used to map histone modifications and TF binding sites (TFBSs) in various B cell populations and to describe the epigenetic changes accompanying B cell development. Recent studies in different systems indicated that the chromatin of cis-REs is in a pre-active state in stem cells and/or early progenitors before the transcriptional initiation, leading to the concept of “gene priming” (85). The priming is thought to be driven by a specific class of TFs called “pioneer TFs,” that are able to induce the early chromatin changes during the gene activation process (86). Pioneer TFs are thought to mark certain loci for downstream activation during development. Cis-RE bookmarking by pioneer TFs was first described in the mouse liver where FoxA and Gata factors were found to bind the liver-specific enhancer of the *alb1* gene in the precursor gut endoderm prior to its activation in nascent liver (87, 88). The appellation “pioneer TF” must meet the following criteria, although these have not always been unambiguously demonstrated in every case: (i) binding to the regulatory region prior to transcription activation, (ii) binding prior to the arrival of other factors, (iii) binding to their target sites in condensed chromatin, and (iv) being able to induce chromatin modifications and/or remodeling in order to render the locus accessible for downstream TFs [reviewed in Ref. (86)]. It is important to mention that in the majority of cases, it is difficult to establish unequivocally the exact binding chronology of a set of TFs at a given locus (see **Figure 4**).

Several studies have described primed enhancers (sometimes also called poised enhancers) in the hematopoietic system (85, 89). Primed enhancers refer to distal regulatory elements that harbor H3K4me1 mark but lack acetylation marks such as H3K27ac and H3K9ac; their associated genes are therefore not transcribed. In contrast, active enhancers harbor both H3K4me1 mark and acetylation marks and their associated genes are transcribed. According to the current priming models, once a cell has reached terminal differentiation, its enhancer repertoire is completely established and maintained by cooperatively acting lineage-specific TFs. Inducible or regulated TFs that are activated by extracellular stimuli operate within this predetermined framework, landing close to where master regulators are already bound (**Figure 5**). However, this model was recently challenged by the identification of a novel class of



**FIGURE 4 | Simplified scheme of stepwise enhancer activation based on the pioneering model.** A pioneer TF is exemplified by the ETS factor PU.1; the transcription start site (TSS) is indicated by a red or green arrow and the enhancer element is schematized by four nucleosomes. For simplicity, the nucleosomes covering the rest of the DNA including the promoter region are not indicated. **(A)** The pioneer TF recognizes and binds its cognate site in condensed chromatin.

**(B)** Recruitment of histone methyltransferases (HMTs) and chromatin remodeling complexes (CRC) which prime the enhancer for subsequent activation. At this step, the enhancer now harbors H3K4me1 mark but still lacks acetylation marks. **(C)** Subsequent collaborative binding of downstream TFs accompanied by histone acetyl transferases (HATs) that catalyze histone acetylation, leading to enhancer activation and gene transcription.



**FIGURE 5 | Stimuli-dependent activation of different classes of enhancers in terminally differentiated cells.** **(A)** In differentiated cells, the majority of stimuli-dependent enhancers are primed, i.e., already bound by pioneer TFs and marked by H3K4me1, but inactive. After stimulation, downstream inducible TFs bind to their cognate sites to

activate target enhancers and gene transcription. **(B)** Latent enhancers are special class of enhancers, that are unbound by pioneer TFs and unmarked by H3K4me1 in unstimulated cells. At this stage, their chromatin structure is not well characterized. TF binding, active chromatin marking, and gene activation occur after the stimulation.

enhancers in macrophages. These cis-REs have been called “latent enhancers” and are not bound by TFs and also lack H3K4me1 and acetylation marks under basal or uninduced conditions. However,

they acquire all these features in response to stimulation (90) (see **Figure 5**). These data suggest that the priming may not be absolutely required for all enhancer elements; however, it cannot

be excluded that upon the stimulation the priming occurs before the activation of target enhancers.

## UNDERSTANDING THE ORDER OF EVENTS

In an effort to investigate how the enhancer repertoire is established and maintained during myeloid and B cell development, Mercer et al. have generated long-term cultures of hematopoietic progenitors by enforcing the expression of the E-protein antagonist Id2, which inhibits E2A activity by preventing its binding to DNA. These progenitors, called Id2-HPC, can be differentiated *in vitro* into myeloid and B cell lineages by switching off Id2 expression, therefore effectively restoring E2A activity. Using this system, H3K4me1 mark was mapped in Id2-HPC cells as well as in myeloid and B cells generated *in vitro* from these artificial precursors. Interestingly, it was found that a substantial fraction of the lymphoid and myeloid enhancers were pre-marked by H3K4me1 (i.e., primed) already in MPPs. Thus, multilineage priming of enhancer elements in hematopoietic progenitors precedes commitment to the lymphoid and myeloid cell lineages (85). Motif analysis showed that PU.1 and Runx motifs were over-represented in H3K4me1 enriched loci in Id2-HPC cells, while enhancers of genes activated after B cell differentiation were enriched in E2A and Ebf1 motifs, in addition to PU.1 motif (85). These findings clearly demonstrate a relationship between cell type-specific binding of TFs and the pattern of H3K4me1 enhancer mark. They also indicate a potential role of PU.1 and Runx factors in priming enhancers in hematopoietic progenitors for subsequent downstream activation.

The correlation between PU.1 binding and presence of H3K4me1 mark was also reported in other hematopoietic lineages, including B cells (30) and macrophages (89). However, it was unclear whether the H3K4me1 modification serves as a beacon to recruit PU.1 and other TFs, or whether these TFs can initiate the deposition of H3K4me1 in hematopoietic progenitors. By expressing a tamoxifen-inducible PU.1/ER fusion protein in PU.1-deficient myeloid progenitors (91), Heinz et al. demonstrated that PU.1 binding can induce H3K4me1 deposition at some loci; yet, many loci were found to lack H3K4me1 despite the binding of PU.1, suggesting that additional factors may be required to write this mark (30). In addition, PU.1 was found to bind to loci that were already marked by H3K4me1; in this case PU.1 was found to be able to initiate nucleosome remodeling (30).

An earlier study has shown that the intronic enhancer of the *Pax5* gene is bound and regulated by PU.1, IRF4, IRF8, and NF- $\kappa$ B (55). Interestingly, the chromatin at this enhancer harbors active marks already in progenitors and is bound by PU.1 and IRF factors before *Pax5* transcription takes place in committed pro-B cells (55). It was also shown that the concerted action of PU.1 and Runx1 primes the activation of both promoter and enhancer elements of the *c-fms* gene in myeloid cell (92, 93). All together, these data clearly indicate the pioneering and priming abilities of the master hematopoietic regulator PU.1. This is consistent with its expression during early hematopoietic cell differentiation from HSCs onward and its dynamic collaborative binding with various TFs.

E2A was also found to alter the H3K4me1 pattern at enhancer elements in B cell progenitors, however it is unclear whether E2A

can directly induce *de novo* H3K4 mono-methylation or only modulate the positioning of nucleosomes already pre-marked by H3K4me1 via nucleosome remodeling mechanisms (24).

Other downstream TFs such as Ebf1 and Pax5 were also found to regulate chromatin structure at cis-REs. For example, Ebf1 plays a role in the demethylation of the *Cd79a* promoter in B cell progenitors (94). Ebf1 is also crucially required for the remodeling and activation of chromatin in the *Pax5* promoter region (55). Pax5 regulates chromatin structure by recruiting chromatin-modifying and remodeling complexes to the Pax5 regulated loci (57). Interestingly, Pax5 fulfills both activation and repression functions; it induces active chromatin at promoters and enhancers of activated target genes, while eliminating active chromatin at the regulatory elements of repressed target genes. Pax5 rapidly induces H3K4 methylation and H3K9 acetylation at enhancers and promoters of activated target genes. Pax5 activation function involves direct interaction with the chromatin-remodeling SWI/SNF-like BAF complex, the histone acetyltransferase CBP, and the PTIP protein, which is known to recruit the MLL-containing H3K4 methyltransferase complex to chromatin (95). The repressing activity of Pax5 is mediated by its ability to recruit the NCoR1 co-repressor complex with its associated HDAC3 enzyme, which is likely responsible for histone deacetylation at some Pax5 repressed loci (57). Pax5 was also found to interact with members of the co-repressor Groucho family, thus leading to repression of target genes (96). An intriguing question is how TFs such as Ikaros and Pax5, having both activation and repression abilities, can distinguish which set of genes must be repressed or activated.

FoxO TFs were also described to have pioneering capacity [reviewed in Ref. (97)] and FoxO1 was found to be able to bind to its cognate sites in condensed chromatin. This binding stably perturbs core histone by de-condensing linker histone-compacted chromatin (98), possibly because the FoxO DBD shares structural similarities with the globular domain of the linker histones H1 and H5 (99, 100). Furthermore, the amino-terminal and carboxy-terminal regions of FoxO1 mediate histone H3 and H4 binding (98). By functioning as pioneer factors, FoxO TFs might open condensed regions and allow the binding of other TFs.

Overall, these studies demonstrated that chromatin structure in hematopoietic progenitors and committed cells can act as a beacon for binding of some TFs. Conversely, TFs such as PU.1, E2A, Ebf1, and many others, can modulate chromatin features at cis-REs to create or enhance a chromatin environment favorable for the binding of additional TFs. The priming and activation of cis-REs requires the collaborative and/or cooperative action of several TFs. For example, in B cells, the pioneer TF PU.1 co-occupies enhancers with E2A, Ebf1, and Oct2, while in macrophages it binds together with AP-1 and C/EBP (30). However, the synergy between pioneer and downstream TFs is not simply hierarchical but also involves cross-regulatory interactions. For example, at certain loci, PU.1 binding in B cells depends on E2A and Ebf1 (30) in spite of the fact that these two factors were not clearly identified as pioneer TFs. What regulates whether PU.1 binds by itself or requires other factors is not known, but is likely to involve the precise binding site and/or the local chromatin structure. It is also not established whether the pioneer TFs identified so far are a special class of factors with unique properties, or whether most factors can



act as pioneers in the right context. Thus, the term “pioneer TF” does not have an absolute meaning, but should rather be viewed as a useful descriptor for properties identified in specific cases. Indeed, a downstream TF can act as a pioneer for an upstream TF, and vice versa, in a context- and locus-dependent manner. Therefore, many TFs involved in the priming of cis-RE can fall into the category of pioneer TFs. However, as mentioned above, it is often difficult to unambiguously monitor the precise chronological binding order of a set of TFs and corresponding epigenetic modifications under *in vivo* conditions. Thus, instead of using the term pioneer when the evidence is scarce, it may be better to rather speak about collaborative action of TFs at a given locus.

## CONCLUDING REMARKS

In summary, the questions of who is on first, the chromatin or the TF, when, and why/how are still largely unanswered. In some physiological situations, specific chromatin features must precede and are required for TF binding, while in other situations the TF binding initiates a series of epigenetic events eventually required for the recruitment of downstream TFs. The extensive efforts that were made to investigate transcriptional and epigenetic regulation of B cells and other hematopoietic lineages identified several mechanisms of cross-regulation between TFs, chromatin modifiers, and the pre-existing chromatin landscape. The interactions between the actors cited above are very likely to be controlled by environmental, spatial, and temporal signals that remain to be defined. Also, many additional factors – TFs, chromatin modifiers, non-coding RNAs . . . – remain to be tested for their potential role in the hematopoietic system or in B cells. However, achieving a deeper understanding of the mechanisms involved will require the ability to examine single cells in real-time to understand how the interplay between chromatin and TFs is orchestrated and unambiguously determine causal relationships. Also, the ability to genetically manipulate the system, not only at the level of the TFs or other trans-acting factors, but also of the cis-REs, e.g., by using the newly developed CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats) system (101) will be invaluable to further our understanding.

## ACKNOWLEDGMENTS

This work was supported by the Novartis Research Foundation and the Swiss National Science Foundation SystemsX Cell Plasticity grant. We thank Benjamin Herquel (Max Planck Institute, Freiburg), Antonius G. Rolink (Basel University DBM, Basel), Makoto Saito, and Roger Clerc for critical reading of the manuscript and valuable comments.

## REFERENCES

- Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* (1997) **389**(6648):251–60. doi:10.1038/38444
- Kooistra SM, Helin K. Molecular mechanisms and potential functions of histone demethylases. *Nat Rev Mol Cell Biol* (2012) **13**(5):297–311. doi:10.1038/nrm3327
- Lei H, Oh SP, Okano M, Jütermann R, Goss KA, Jaenisch R, et al. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* (1996) **122**(10):3195–205.
- Bird AP, Wolffe AP. Methylation-induced repression – belts, braces, and chromatin. *Cell* (1999) **99**(5):451–4. doi:10.1016/S0092-8674(00)81532-9
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* (2009) **462**(7271):315–22. doi:10.1038/nature08514
- Yoder JA, Soman NS, Verdine GL, Bestor TH. DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *J Mol Biol* (1997) **270**(3):385–95. doi:10.1006/jmbi.1997.1125
- Okano M, Takebayashi S, Okumura K, Li E. Assignment of cytosine-5 DNA methyltransferases Dnmt3a and Dnmt3b to mouse chromosome bands 12A2-A3 and 2H1 by in situ hybridization. *Cytogenet Cell Genet* (1999) **86**(3–4):333–4. doi:10.1159/000015331
- Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol* (2010) **11**(9):607–20. doi:10.1038/nrm2950
- Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* (2010) **466**(7310):1129–33. doi:10.1038/nature09303
- Zhang P, Su L, Wang Z, Zhang S, Guan J, Chen Y, et al. The involvement of 5-hydroxymethylcytosine in active DNA demethylation in mice. *Biol Reprod* (2012) **86**(4):104. doi:10.1095/biolreprod.111.096073
- Mueller-Planitz F, Klinker H, Becker PB. Nucleosome sliding mechanisms: new twists in a looped history. *Nat Struct Mol Biol* (2013) **20**(9):1026–32. doi:10.1038/nsmb.2648
- Watanabe S, Peterson CL. The INO80 family of chromatin-remodeling enzymes: regulators of histone variant dynamics. *Cold Spring Harb Symp Quant Biol* (2010) **75**:35–42. doi:10.1101/sqb.2010.75.063
- Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet* (2009) **10**(3):155–9. doi:10.1038/nrg2521
- Iwasaki H, Akashi K. Myeloid lineage commitment from the hematopoietic stem cell. *Immunity* (2007) **26**(6):726–40. doi:10.1016/j.immuni.2007.06.004
- Kondo M, Weissman IL, Akashi K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* (1997) **91**(5):661–72. doi:10.1016/S0092-8674(00)80453-5
- Akashi K, Traver D, Miyamoto T, Weissman IL. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* (2000) **404**(6774):193–7. doi:10.1038/35004599
- Adolfsson J, Borge OJ, Bryder D, Theilgaard-Mönch K, Astrand-Grundström I, Sitnicka E, et al. Upregulation of Flt3 expression within the bone marrow Lin(–)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* (2001) **15**(4):659–69. doi:10.1016/S1074-7613(01)00220-5
- Adolfsson J, Månsson R, Buza-Vidas N, Hultquist A, Liuba K, Jensen CT, et al. Identification of Flt3+ lympho-myeloid stem cells lacking erythromegakaryocytic potential: a revised road map for adult blood lineage commitment. *Cell* (2005) **121**(2):295–306. doi:10.1016/j.cell.2005.02.013
- Igarashi H, Gregory SC, Yokota T, Sakaguchi N, Kincade PW. Transcription from the RAG1 locus marks the earliest lymphocyte progenitors in bone marrow. *Immunity* (2002) **17**(2):117–30. doi:10.1016/S1074-7613(02)00366-7
- Inlay MA, Bhattacharya D, Sahoo D, Serwold T, Seita J, Karsunky H, et al. Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes Dev* (2009) **23**(20):2376–81. doi:10.1101/gad.1836009
- Matthias P, Rolink AG. Transcriptional networks in developing and mature B cells. *Nat Rev Immunol* (2005) **5**(6):497–508. doi:10.1038/nri1633
- Hendriks RW, Middendorp S. The pre-BCR checkpoint as a cell-autonomous proliferation switch. *Trends Immunol* (2004) **25**(5):249–56. doi:10.1016/j.it.2004.02.011
- Perlot T, Alt FW. Cis-regulatory elements and epigenetic changes control genomic rearrangements of the IgH locus. *Adv Immunol* (2008) **99**:1–32. doi:10.1016/S0065-2776(08)00601-9
- Lin YC, Jhunjhunwala S, Benner C, Heinz S, Welinder E, Månsson R, et al. A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat Immunol* (2010) **11**(7):635–43. doi:10.1038/ni.1891
- Zhang Q, Iida R, Yokota T, Kincade PW. Early events in lymphopoiesis: an update. *Curr Opin Hematol* (2013) **20**(4):265–72. doi:10.1097/MOH.0b013e3283612628
- Wang F, Tong Q. Transcription factor PU.1 is expressed in white adipose and inhibits adipocyte differentiation. *Am J Physiol Cell Physiol* (2008) **295**(1):C213–20. doi:10.1152/ajpcell.00422.2007



27. McKercher SR, Torbett BE, Anderson KL, Henkel GW, Vestal DJ, Baribault H, et al. Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. *EMBO J* (1996) **15**(20):5647–58.
28. Iwasaki H, Somoza C, Shigematsu H, Duprez EA, Iwasaki-Arai J, Mizuno S, et al. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood* (2005) **106**(5):1590–600. doi:10.1182/blood-2005-03-0860
29. Klemsz MJ, McKercher SR, Celada A, Van Beveren C, Maki RA. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. *Cell* (1990) **61**(1):113–24. doi:10.1016/0092-8674(90)90219-5
30. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* (2010) **38**(4):576–89. doi:10.1016/j.molcel.2010.05.004
31. Back J, Allman D, Chan S, Kastner P. Visualizing PU.1 activity during hematopoiesis. *Exp Hematol* (2005) **33**(4):395–402. doi:10.1016/j.exphem.2004.12.010
32. DeKoter RP, Singh H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science* (2000) **288**(5470):1439–41. doi:10.1126/science.288.5470.1439
33. Georgopoulos K, Bigby M, Wang JH, Molnar A, Wu P, Winandy S, et al. The Ikaros gene is required for the development of all lymphoid lineages. *Cell* (1994) **79**(1):143–56. doi:10.1016/0092-8674(94)90407-3
34. Yoshida T, Ng SY, Zuniga-Pflucker JC, Georgopoulos K. Early hematopoietic lineage restrictions directed by Ikaros. *Nat Immunol* (2006) **7**(4):382–91. doi:10.1038/ni1314
35. Ng SY, Yoshida T, Zhang J, Georgopoulos K. Genome-wide lineage-specific transcriptional networks underscore Ikaros-dependent lymphoid priming in hematopoietic stem cells. *Immunity* (2009) **30**(4):493–507. doi:10.1016/j.immuni.2009.01.014
36. Reynaud D, Demarco IA, Reddy KL, Schjerven H, Bertolino E, Chen Z, et al. Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat Immunol* (2008) **9**(8):927–36. doi:10.1038/ni.1626
37. Heizmann B, Kastner P, Chan S. Ikaros is absolutely required for pre-B cell differentiation by attenuating IL-7 signals. *J Exp Med* (2013) **210**(13):2823–32. doi:10.1084/jem.20131735
38. Sridharan R, Smale ST. Predominant interaction of both Ikaros and Helios with the NuRD complex in immature thymocytes. *J Biol Chem* (2007) **282**(41):30227–38. doi:10.1074/jbc.M702541200
39. O'Neill DW, Schoetz SS, Lopez RA, Castle M, Rabinowitz L, Shor E, et al. An ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. *Mol Cell Biol* (2000) **20**(20):7572–82. doi:10.1128/MCB.20.20.7572-7582.2000
40. Kim J, Sif S, Jones B, Jackson A, Koipally J, Heller E, et al. Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. *Immunity* (1999) **10**(3):345–55. doi:10.1016/S1074-7613(00)80034-5
41. Cobb BS, Morales-Alcay S, Kleiger G, Brown KE, Fisher AG, Smale ST. Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes Dev* (2000) **14**(17):2146–60. doi:10.1101/gad.816400
42. Brown KE, Guest SS, Smale ST, Hahn K, Merkenschlager M, Fisher AG. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* (1997) **91**(6):845–54. doi:10.1016/S0092-8674(00)80472-9
43. Bain G, Maandag EC, Izon DJ, Amsen D, Kruisbeek AM, Weintraub BC, et al. E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell* (1994) **79**(5):885–92. doi:10.1016/0092-8674(94)90077-9
44. Zhuang Y, Soriano P, Weintraub H. The helix-loop-helix gene E2A is required for B cell formation. *Cell* (1994) **79**(5):875–84. doi:10.1016/0092-8674(94)90076-0
45. Kee BL, Murre C. Induction of early B cell factor (EBF) and multiple B lineage genes by the basic helix-loop-helix transcription factor E12. *J Exp Med* (1998) **188**(4):699–713. doi:10.1084/jem.188.4.699
46. Welinder E, Mansson R, Mercer EM, Bryder D, Sigvardsson M, Murre C. The transcription factors E2A and HEB act in concert to induce the expression of FOXO1 in the common lymphoid progenitor. *Proc Natl Acad Sci U S A* (2011) **108**(42):17402–7. doi:10.1073/pnas.1111766108
47. Daburon V, Mella S, Plouhinec JL, Mazan S, Crozatier M, Vincent A. The meta-zoan history of the COE transcription factors. Selection of a variant HLH motif by mandatory inclusion of a duplicated exon in vertebrates. *BMC Evol Biol* (2008) **8**:131. doi:10.1186/1471-2148-8-131
48. Hagman J, Gutch MJ, Lin H, Grosschedl R. EBF contains a novel zinc coordination motif and multiple dimerization and transcriptional activation domains. *EMBO J* (1995) **14**(12):2907–16.
49. Zandi S, Ahsberg J, Tsapogas P, Stjernberg J, Qian H, Sigvardsson M. Single-cell analysis of early B-lymphocyte development suggests independent regulation of lineage specification and commitment in vivo. *Proc Natl Acad Sci U S A* (2012) **109**(39):15871–6. doi:10.1073/pnas.1210144109
50. Lin H, Grosschedl R. Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature* (1995) **376**(6537):263–7. doi:10.1038/376263a0
51. Zandi S, Mansson R, Tsapogas P, Zetterblad J, Bryder D, Sigvardsson M. EBF1 is essential for B-lineage priming and establishment of a transcription factor network in common lymphoid progenitors. *J Immunol* (2008) **181**(5):3364–72.
52. Treiber T, Mandel EM, Pott S, Györy I, Firner S, Liu ET, et al. Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription-independent poisoning of chromatin. *Immunity* (2010) **32**(5):714–25. doi:10.1016/j.immuni.2010.04.013
53. Roessler S, Györy I, Imhof S, Spivakov M, Williams RR, Busslinger M, et al. Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. *Mol Cell Biol* (2007) **27**(2):579–94. doi:10.1128/MCB.01192-06
54. Thal MA, Carvalho TL, He T, Kim HG, Gao H, Hagman J, et al. Ebf1-mediated down-regulation of Id2 and Id3 is essential for specification of the B cell lineage. *Proc Natl Acad Sci U S A* (2009) **106**(2):552–7. doi:10.1073/pnas.0802550106
55. Decker T, Pasca di Magliano M, McManus S, Sun Q, Bonifer C, Tagoh H, et al. Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. *Immunity* (2009) **30**(4):508–20. doi:10.1016/j.immuni.2009.01.012
56. Urbánek P, Wang ZQ, Fetka I, Wagner EF, Busslinger M. Complete block of early B cell differentiation and altered patterning of the posterior mid-brain in mice lacking Pax5/BSAP. *Cell* (1994) **79**(5):901–12. doi:10.1016/0092-8674(94)90079-5
57. McManus S, Ebert A, Salvagiotto G, Medvedovic J, Sun Q, Tamir I, et al. The transcription factor Pax5 regulates its target genes by recruiting chromatin-modifying proteins in committed B cells. *EMBO J* (2011) **30**(12):2388–404. doi:10.1038/emboj.2011.140
58. Cobaleda C, Jochum W, Busslinger M. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature* (2007) **449**(7161):473–7. doi:10.1038/nature06159
59. Dengler HS, Baracho GV, Omori SA, Bruckner S, Arden KC, Castrillon DH, et al. Distinct functions for the transcription factor Foxo1 at various stages of B cell differentiation. *Nat Immunol* (2008) **9**(12):1388–98. doi:10.1038/ni.1667
60. Amin RH, Schlissel MS. Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. *Nat Immunol* (2008) **9**(6):613–22. doi:10.1038/ni.1612
61. Fahl SP, Crittenden RB, Allman D, Bender TP. c-Myb is required for pro-B cell differentiation. *J Immunol* (2009) **183**(9):5582–92. doi:10.4049/jimmunol.0901187
62. Seo W, Ikawa T, Kawamoto H, Taniuchi I. Runx1-Cbfbeta facilitates early B lymphocyte development by regulating expression of Ebf1. *J Exp Med* (2012) **209**(7):1255–62. doi:10.1084/jem.20112745
63. Huang G, Zhang P, Hirai H, Elf S, Yan X, Chen Z, et al. PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. *Nat Genet* (2008) **40**(1):51–60. doi:10.1038/ng0208-255a
64. Mak KS, Funnell AP, Pearson RC, Crossley M. PU.1 and haematopoietic cell fate: dosage matters. *Int J Cell Biol* (2011) **2011**:808524. doi:10.1155/2011/808524
65. Aloia L, Di Stefano B, Di Croce L. Polycomb complexes in stem cells and embryonic development. *Development* (2013) **140**(12):2525–34. doi:10.1242/dev.091553
66. Su IH, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT, et al. Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat Immunol* (2003) **4**(2):124–31. doi:10.1038/ni876
67. Aldiri I, Vetter ML. PRC2 during vertebrate organogenesis: a complex in transition. *Dev Biol* (2012) **367**(2):91–9. doi:10.1016/j.ydbio.2012.04.030

68. van der Lugt NM, Domen J, Linders K, van Roon M, Robanus-Maandag E, te Riele H, et al. Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the bmi-1 proto-oncogene. *Genes Dev* (1994) 8(7):757–69. doi:10.1101/gad.8.7.757
69. Akasaka T, Tsuji K, Kawahira H, Kanno M, Harigaya K, Hu L, et al. The role of mel-18, a mammalian Polycomb group gene, during IL-7-dependent proliferation of lymphocyte precursors. *Immunity* (1997) 7(1):135–46. doi:10.1016/S1074-7613(00)80516-6
70. Coré N, Bel S, Gaunt SJ, Aurrand-Lions M, Pearce J, Fisher A, et al. Altered cellular proliferation and mesoderm patterning in Polycomb-M33-deficient mice. *Development* (1997) 124(3):721–9.
71. Park IK, Qian D, Kiel M, Becker MW, Pihajla M, Weissman IL, et al. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* (2003) 423(6937):302–5. doi:10.1038/nature01587
72. Oguro H, Iwama A, Morita Y, Kamijo T, van Lohuizen M, Nakauchi H. Differential impact of Ink4a and Arf on hematopoietic stem cells and their bone marrow microenvironment in Bmi1-deficient mice. *J Exp Med* (2006) 203(10):2247–53. doi:10.1084/jem.20052477
73. Oguro H, Yuan J, Ichikawa H, Ikawa T, Yamazaki S, Kawamoto H, et al. Poised lineage specification in multipotential hematopoietic stem and progenitor cells by the polycomb protein Bmi1. *Cell Stem Cell* (2010) 6(3):279–86. doi:10.1016/j.stem.2010.01.005
74. Calés C, Román-Trufero M, Pavón L, Serrano I, Melgar T, Endoh M, et al. Inactivation of the polycomb group protein Ring1B unveils an antiproliferative role in hematopoietic cell expansion and cooperation with tumorigenesis associated with Ink4a deletion. *Mol Cell Biol* (2008) 28(3):1018–28. doi:10.1128/MCB.01136-07
75. Mochizuki-Kashio M, Mishima Y, Miyagi S, Negishi M, Saraya A, Konuma T, et al. Dependency on the polycomb gene Ezh2 distinguishes fetal from adult hematopoietic stem cells. *Blood* (2011) 118(25):6553–61. doi:10.1182/blood-2011-03-340554
76. Yamaguchi T, Cubizolles F, Zhang Y, Reichert N, Kohler H, Seiser C, et al. Histone deacetylases 1 and 2 act in concert to promote the G1-to-S progression. *Genes Dev* (2010) 24(5):455–69. doi:10.1101/gad.552310
77. Trowbridge JJ, Snow JW, Kim J, Orkin SH. DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell Stem Cell* (2009) 5(4):442–9. doi:10.1016/j.stem.2009.08.016
78. Bröske AM, Vockentanz L, Kharazi S, Huska MR, Mancini E, Scheller M, et al. DNA methylation protects hematopoietic stem cell multipotency from myeloid erythroid restriction. *Nat Genet* (2009) 41(11):1207–15. doi:10.1038/ng.463
79. Challen GA, Sun D, Jeong M, Luo M, Jelinek J, Berg JS, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet* (2012) 44(1):23–31. doi:10.1038/ng.1009
80. Nakase H, Takahama Y, Akamatsu Y. Effect of CpG methylation on RAG1/RAG2 reactivity: implications of direct and indirect mechanisms for controlling V(D)J cleavage. *EMBO Rep* (2003) 4(8):774–80. doi:10.1038/sj.embor.embor904
81. Selimyan R, Gerstein RM, Ivanova I, Precht P, Subrahmanyam R, Perlot T, et al. Localized DNA demethylation at recombination intermediates during immunoglobulin heavy chain gene assembly. *PLoS Biol* (2013) 11(1):e1001475. doi:10.1371/journal.pbio.1001475
82. Cherry SR, Beard C, Jaenisch R, Baltimore D. V(D)J recombination is not activated by demethylation of the kappa locus. *Proc Natl Acad Sci U S A* (2000) 97(15):8467–72. doi:10.1073/pnas.150218497
83. Gray KS, Forrest JC, Speck SH. The de novo methyltransferases DNMT3a and DNMT3b target the murine gammaherpesvirus immediate-early gene 50 promoter during establishment of latency. *J Virol* (2010) 84(10):4946–59. doi:10.1128/JVI.00060-10
84. Rickert RC, Roes J, Rajewsky K. B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucleic Acids Res* (1997) 25(6):1317–8. doi:10.1093/nar/25.6.1317
85. Mercer EM, Lin YC, Benner C, Jhunjhunwala S, Dutkowski J, Flores M, et al. Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors. *Immunity* (2011) 35(3):413–25. doi:10.1016/j.immuni.2011.06.013
86. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* (2011) 25(21):2227–41. doi:10.1101/gad.176826.111
87. Gualdi R, Bossard P, Zheng M, Hamada Y, Coleman JR, Zaret KS. Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. *Genes Dev* (1996) 10(13):1670–82. doi:10.1101/gad.10.13.1670
88. Bossard P, Zaret KS. GATA transcription factors as potentiators of gut endoderm differentiation. *Development* (1998) 125(24):4909–17.
89. Ghisletti S, Barozzi I, Mietton F, Polletti S, De Santa F, Venturini E, et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* (2010) 32(3):317–28. doi:10.1016/j.immuni.2010.02.008
90. Ostuni R, Piccolo V, Barozzi I, Polletti S, Termanini A, Bonifacio S, et al. Latent enhancers activated by stimulation in differentiated cells. *Cell* (2013) 152(1–2):157–71. doi:10.1016/j.cell.2012.12.018
91. Walsh JC, DeKoter RP, Lee HJ, Smith ED, Lancki DW, Gurish MF, et al. Cooperative and antagonistic interplay between PU.1 and GATA-2 in the specification of myeloid cell fates. *Immunity* (2002) 17(5):665–76. doi:10.1016/S1074-7613(02)00452-1
92. Kryszinska H, Hoogenkamp M, Ingram R, Wilson N, Tagoh H, Laslo P, et al. A two-step, PU.1-dependent mechanism for developmentally regulated chromatin remodeling and transcription of the c-fms gene. *Mol Cell Biol* (2007) 27(3):878–87. doi:10.1128/MCB.01915-06
93. Hoogenkamp M, Lichtinger M, Kryszinska H, Lancrin C, Clarke D, Williamson A, et al. Early chromatin unfolding by RUNX1: a molecular explanation for differential requirements during specification versus maintenance of the hematopoietic gene expression program. *Blood* (2009) 114(2):299–309. doi:10.1182/blood-2008-11-191890
94. Maier H, Ostraat R, Gao H, Fields S, Shinton SA, Medina KL, et al. Early B cell factor cooperates with Runx1 and mediates epigenetic changes associated with mb-1 transcription. *Nat Immunol* (2004) 5(10):1069–77. doi:10.1038/ni1119
95. Cho YW, Hong T, Hong S, Guo H, Yu H, Kim D, et al. PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. *J Biol Chem* (2007) 282(28):20395–406. doi:10.1074/jbc.M701574200
96. Eberhard D, Jiménez G, Heavey B, Busslinger M. Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. *EMBO J* (2000) 19(10):2292–303. doi:10.1093/emboj/19.10.2292
97. Lalmansingh AS, Karmakar S, Jin Y, Nagaich AK. Multiple modes of chromatin remodeling by Forkhead box proteins. *Biochim Biophys Acta* (2012) 1819(7):707–15. doi:10.1016/j.bbagr.2012.02.018
98. Hattar M, Cirillo LA. Chromatin opening and stable perturbation of core histone:DNA contacts by FoxO1. *J Biol Chem* (2007) 282(49):35583–93. doi:10.1074/jbc.M704735200
99. Clark KL, Halay ED, Lai E, Burley SK. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* (1993) 364(6436):412–20. doi:10.1038/364412a0
100. Ramakrishnan V, Finch JT, Graziano V, Lee PL, Sweet RM. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* (1993) 362(6417):219–23. doi:10.1038/362219a0
101. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* (2013) 8(11):2281–308. doi:10.1038/nprot.2013.143

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 January 2014; accepted: 25 March 2014; published online: 11 April 2014.  
Citation: Choukrallah MA and Matthias P (2014) The interplay between chromatin and transcription factor networks during B cell development: who pulls the trigger first? *Front. Immunol.* 5:156. doi: 10.3389/fimmu.2014.00156

This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Choukrallah and Matthias. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The Bright side of hematopoiesis: regulatory roles of ARID3a/Bright in human and mouse hematopoiesis

Michelle L. Ratliff<sup>1</sup>, Troy D. Templeton<sup>2</sup>, Julie M. Ward<sup>3</sup> and Carol F. Webb<sup>1,2,3 \*</sup>

<sup>1</sup> Immunobiology and Cancer Research, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA

<sup>2</sup> Department of Cell Biology, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

<sup>3</sup> Department of Microbiology and Immunology, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

Shiv Pillai, Harvard Medical School, USA

Rodney P. DeKoter, The University of Western Ontario, Canada

## \*Correspondence:

Carol F. Webb, Immunobiology and Cancer Research Program, MS 29, Oklahoma Medical Research Foundation, 825 N.E. 13th Street, Oklahoma City, OK 73104, USA  
e-mail: carol-webb@omrf.org

ARID3a/Bright is a DNA-binding protein that was originally discovered for its ability to increase immunoglobulin transcription in antigen-activated B cells. It interacts with DNA as a dimer through its ARID, or A/T-rich interacting domain. In association with other proteins, ARID3a increased transcription of the immunoglobulin heavy chain and led to improved chromatin accessibility of the heavy chain enhancer. Constitutive expression of ARID3a in B lineage cells resulted in autoantibody production, suggesting its regulation is important. Abnormal ARID3a expression has also been associated with increased proliferative capacity and malignancy. Roles for ARID3a in addition to interactions with the immunoglobulin locus were suggested by transgenic and knockout mouse models. Over-expression of ARID3a resulted in skewing of mature B cell subsets and altered gene expression patterns of follicular B cells, whereas loss of function resulted in loss of B1 lineage B cells and defects in hematopoiesis. More recent studies showed that loss of ARID3a in adult somatic cells promoted developmental plasticity, alterations in gene expression patterns, and lineage fate decisions. Together, these data suggest new regulatory roles for ARID3a. The genes influenced by ARID3a are likely to play pivotal roles in lineage decisions, highlighting the importance of this understudied transcription factor.

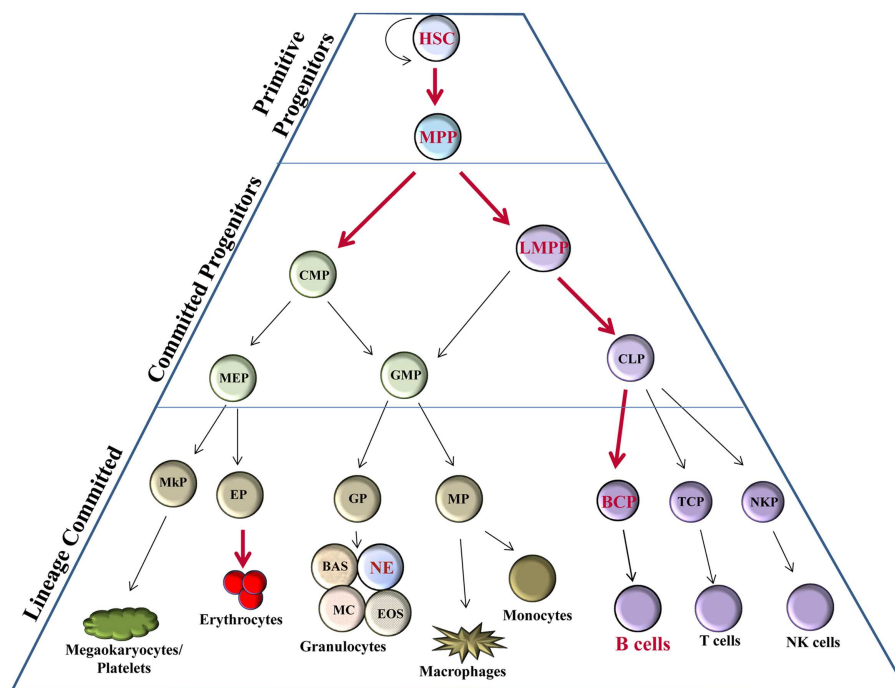
**Keywords:** Bright, ARID3a, hematopoietic regulation, B cell development, gene regulation

Bright (B cell regulator of immunoglobulin heavy chain transcription) is a 70 kDa DNA-binding protein first characterized in the mouse as a component of a protein complex associated with increased transcription of the immunoglobulin heavy chain (IgH) locus in activated B lymphocytes (1–3). *Bright*, also known as *DRIL1*, *E2FBP1*, or *ARID3a* (the designation for the human ortholog, hereafter referred to as Bright), is a member of the A + T rich interaction domain (ARID) protein family, many of which have been shown recently to have epigenetic regulatory functions [reviewed in Ref. (4–6)]. These proteins bind to A + T rich DNA sequences and are typically members of larger chromatin modulatory complexes. Bright, and other ARID3 family members, require dimerization for DNA-binding activity and contain an extended DNA-binding domain that confers increased DNA sequence specificity to these proteins compared to other ARID family members (7–9). Although Bright was the first member of this family identified in mammalian cells, its functions have only begun to be elucidated. Previously, Bright expression in adult, mouse, and human cells was thought to be largely restricted to B lymphocyte lineage cells. However, more recently, we and others have shown that Bright plays important regulatory functions in early hematopoiesis. Although Bright expression is restricted in adults, it is more widely expressed in the embryo/fetus and plays important regulatory roles in embryonic stem cell differentiation (10). These data also highlight novel roles for Bright in gene repression. This article will emphasize the regulatory roles of Bright in

hematopoiesis and will summarize new contributions pertaining to its regulatory capacity in those and other cell types.

## BRIGHT AND HSCs

From a historical perspective, the majority of studies involving Bright have aimed at understanding its roles in B lymphocytes. However, recent evidence suggests Bright may play an even broader role in the development of hematopoietic lineage cells. Hematopoietic stem cells (HSCs) have the capacity to self-renew or to differentiate into other precursors that will eventually produce all mature blood cell types. Differentiation of hematopoietic progenitors occurs primarily along three pathways: erythroid, myeloid, and lymphoid lineages (Figure 1). An intricate network of transcription factors contribute to HSC fate decisions, with more than 20 transcription factors implicated in the development of various hematopoietic subpopulations [reviewed in Ref. (11)], such as growth factor independence 1 (Gfi1), E2A, and Ikaros family zinc-finger protein 1 (Ikaros) in lymphoid lineage regulation; CCAAT-enhancer binding protein alpha (C/EBP $\alpha$ ), GATA1, and PU.1 for myeloid lineage decisions [reviewed in Ref. (11–13)]. Bright is expressed in HSCs in both mouse and man [(14–16); and our unpublished data] and appears to be required for development of several early progenitor subsets including multipotent progenitors (MPPs) and lymphoid-primed MPP (LMPP) (Figure 1). Therefore, Bright contributes to early progenitor ontogeny, which may ultimately affect the development of multiple lineages.



**FIGURE 1 | Bright expression in hematopoiesis.** Hematopoietic progenitor populations [as described (11)] indicate stages, which express Bright/ARID3a (red font) versus those not known to express Bright (black font). Thick red arrows indicate stages of developmental progression for which Bright function is important. HSC, hematopoietic stem cell; MPP, multipotent progenitor; LMPP, lymphoid-primed multipotent progenitor; CMP, common

myeloid progenitor; CLP, common lymphoid progenitor; GMP, granulocyte-macrophage progenitor; MEP, megakaryocyte-erythrocyte progenitor; BCP, B cell progenitor; TCP, T cell progenitor; NKP, NK cell progenitor; MP, macrophage progenitor; GP, granulocyte progenitor; EP, erythrocyte progenitor; MkP, megakaryocyte progenitor; BAS, basophils; NE, neutrophils; MC, mast cells; EOS, eosinophils.

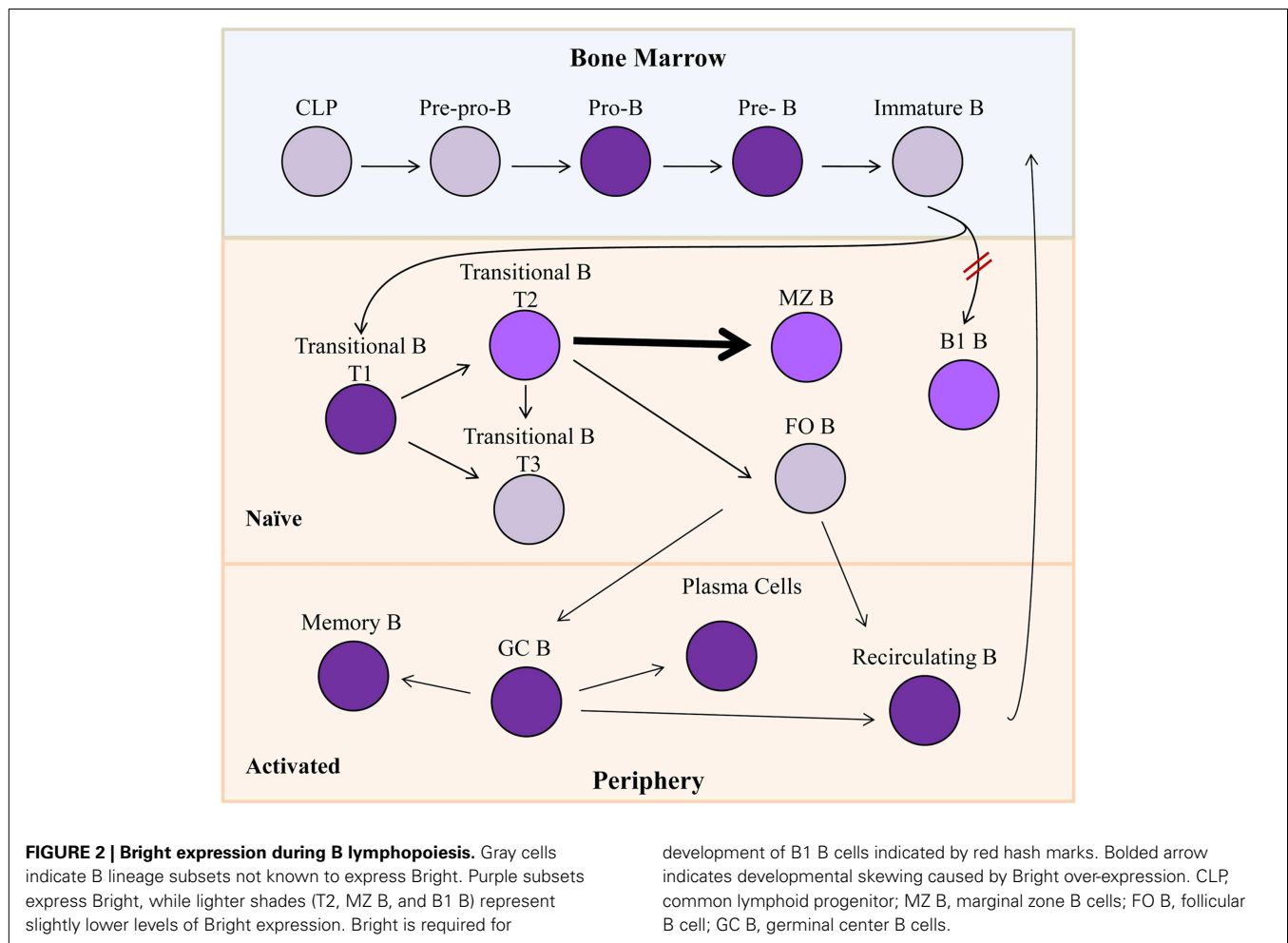
Bright knockout mice die between E11.5 and E13.5 as a result of defects in erythroid lineage differentiation (16). Bright knockout embryos have severe pallor and show fewer mature erythrocytes by flow cytometry. Embryonic death in Bright-deficient mice coincides with the shift from primitive hematopoiesis in the yolk sac to definitive hematopoiesis in the fetal liver. Numbers of fetal liver  $\text{lin}^- \text{cKit}^{\text{hi}} \text{Sca1}^+ \text{CD150}^+ \text{CD48}^-$  HSCs in these embryos were reduced by >90%, while LSKs ( $\text{Lin}^- \text{Sca1}^+ \text{cKit}^+$  cells that include HSC and MPP populations) were decreased in Bright deficient versus wild-type littermate controls by 80% (16). Numbers of common myeloid progenitors (CMPs) and common lymphoid progenitors (CLPs) were also decreased in Bright knockout fetal livers, but to a lesser degree. This suggests that Bright may have greater effects in HSCs than in later precursor subsets. Bright knockout fetal liver cells were also impaired in their ability to generate erythroblast, erythromyeloid, and B lymphocyte colonies in *in vitro* methylcellulose cultures compared to wild-type cells (16). Therefore, Bright may also contribute to the expansion and development of these hematopoietic cells. Importantly, these data confirm the importance of Bright for normal erythroid differentiation.

Rare Bright knockout mice (<1%) survived to adulthood for unknown reasons. These adult Bright-deficient mice exhibited reduced numbers of hematopoietic precursors in the bone marrow, including LSK, CMP, and CLP subsets, but to a lesser degree than was observed in Bright null embryos (16). Although adult

Bright-deficient mice showed decreased numbers of several early B cell subsets, interestingly, no deficiencies in T cell or erythrocyte development were observed (16). The importance of Bright in other hematopoietic lineages has not been fully explored; however, recent data from The Immunological Genome Project suggest Bright transcripts are expressed in mouse neutrophils where its function is unclear (17). Currently, our data from knockout mouse models suggest that the function of Bright in adult bone marrow may be more critical for the development of B lymphocytes and early precursor subpopulations of those cells than for other hematopoietic cell types.

### BRIGHT SIDE OF B CELLS

Common lymphoid progenitors give rise to early subpopulations of precursor B cells (Figure 2), which eventually develop through multiple differentiation states in both the bone marrow and the periphery to result in fully differentiated, mature B cells. Pro-B cells express RAG1/2 gene products, and rearrange IgH gene segments (18, 19), and in man, they are the first B cell progenitors to transcribe Bright (14). Bright expression occurs in both mouse and human subsets of pre-B cells, transitional B cells, activated and memory B cells, and plasma cells (14, 15) (Figure 2). However, the majority of resting, naïve, and mature, peripheral B lymphocytes do not express Bright as evidenced by the absence of Bright transcripts in the majority of circulating blood cells in man, and in follicular spleen cells in the mouse (14, 15). Both human and



mouse innate-like B cells, represented by peritoneal cavity B1 B lymphocytes and splenic marginal zone (MZ) B lymphocytes in mice, and by B1-like and MZ-like cells in human peripheral blood [reviewed in Ref. (20, 21)], express low levels of Bright [(22); unpublished data]. Thus, Bright expression is tightly regulated at the level of transcription throughout B cell differentiation.

Transgenic mice expressing a dominant negative form of Bright from the CD19 pan-B cell promoter were generated by introduction of point mutations in the DNA-binding domain to test the function of Bright within B lineage cells (8, 22). These transgenic mice had slightly decreased numbers of mature B cells, decreased serum IgM levels, and defective responses against *Streptococci* (22), phenotypes similar to those observed in Bruton's tyrosine kinase (Btk) deficient mice [(23), reviewed in Ref. (24, 25)]. Although these mice developed mature B1 B cells, a major source of IgM in the mouse, those cells were functionally defective in their ability to secrete immunoglobulin (22). Furthermore, the few total Bright knockout mice that survived to adulthood lacked B1 B cells (16), suggesting Bright is important both for development and function of those mature B cells. Adult Bright-deficient mice also showed impaired T cell-dependent IgG<sub>1</sub> production due to defects in class switch recombination (16). Thus, Bright contributes to select functions in mature B cells.

Forced expression of native Bright throughout the B cell lineage suggested that regulation of Bright is critical for normal development of MZ and follicular B cells. Transgenic FVB/N mice constitutively over-expressing Bright from the CD19 promoter exhibited significant increase in immature transitional B and MZ B cells relative to the other splenic B cell populations (26). On the C57BL/6 background, over-expression of Bright also resulted in decreased numbers of follicular (FO) B cells (27). The FO B cells that were present in those mice displayed transcript levels of genes previously shown to be differentially expressed in FO versus MZ B cells (28) that were similar to those observed in MZ B cells. Chimeras generated from mixtures of transgenic and wild-type bone marrow cells also showed preferential development of MZ versus FO B cells (27), suggesting that constitutive Bright expression contributes preferentially to MZ versus FO B cell development.

Bright expression can be induced in mature resting B cells through a number of activating signals, including stimulation with lipopolysaccharide (LPS), CD40 ligand, interleukin-5 (IL-5) plus specific antigen, and with agonistic monoclonal antibodies against CD38 or RP105 (2, 3, 15). Additionally, Epstein-Barr Virus (EBV) infection also activates Bright expression in human B cells (14). Although the function of Bright in B lymphocytes was thought to



be exerted primarily in the nucleus via interactions with A + T-rich DNA sequences, a very small percentage of Bright was discovered to be palmitoylated and diverted to lipid rafts (29). *In vitro* studies demonstrated Bright localization to these rafts increased the signaling threshold of the B cell receptor (BCR). Upon effective activation, Bright was released from the lipid rafts via SUMOylation. Interestingly, Btk remained unphosphorylated when Bright was present, delineating a putative role for Bright in BCR signaling (29). However, FO B cells from transgenic mice with forced expression of Bright did not have elevated levels of Bright in lipid rafts, in contrary to transitional and MZ B cell populations, nor did they have alterations in their ability to flux calcium through the BCR. Together, these data suggested other cell type-specific factors may contribute to Bright-mediated effects through the BCR (27).

### GENE TARGETS FOR BRIGHT

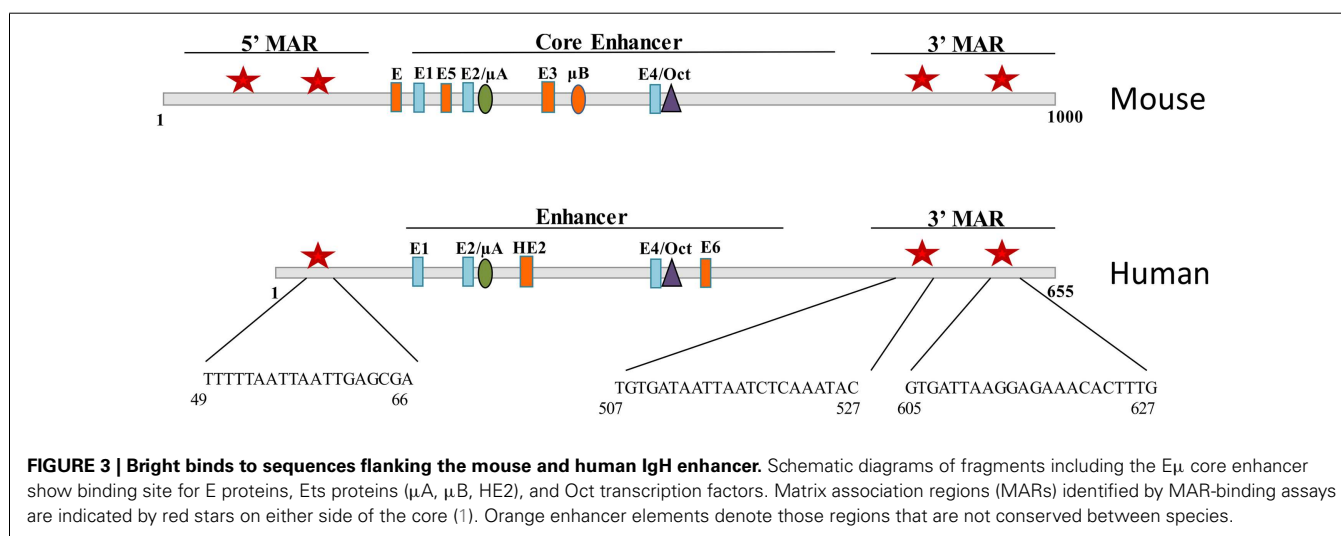
It was originally observed that stimulation of an antigen-specific mouse B cell line with antigen and IL-5 resulted in an increase in immunoglobulin ( $\mu$ ) heavy chain transcription. Further analyses identified two discrete A + T rich elements within the V1 S107 variable heavy chain promoter were bound by a protein complex later identified to contain Bright (3). Intriguingly, further analyses showed potential Bright-binding motifs in about half of the murine  $V_H$  promoters, and binding sites were not restricted to specific V region families (30, 31). Similarly, only a subset of human  $V_H$  promoters had binding sites for Bright (32). These data suggest that Bright may preferentially affect transcription of a subset of IgH genes.

In the mouse, the intronic enhancer of the IgH gene is flanked by 5' and 3' A + T rich regions called matrix associated regions (MARs) that act to tether DNA to the nuclear matrix. Promoter binding sites for Bright were shown to have MAR activity (1), and conversely, Bright was shown to bind to the MARs of the intronic IgH enhancer in the mouse (7). Mouse and human IgH enhancers are similar, but not identical, showing some differences in binding sites for several transcription factors (Figure 3). Bright-binding sites flank both the mouse and human enhancer core sequences. Although Bright did not directly activate transcription of the IgH

enhancer regions *in vitro*, in contrast to promoter fragments that bound Bright (33), MAR regions bound by Bright have been proposed to play important roles establishing chromatin domains important for expression. Indeed, the intronic enhancer region in the mouse locus has been shown to be important for chromatin remodeling, and binding of Bright to sites flanking that region correlated with increased enhancer accessibility (34). Bright can also form tetramers and was found to enhance DNA-bending, suggesting it may also contribute to higher order structures linking the intronic enhancer to specific  $V_H$  region promoters (35). Therefore, we speculate that Bright, like other members of the ARID family, may contribute to epigenetic regulation of the IgH locus.

Recently, additional gene targets for Bright have been identified. Bright was shown to bind to the core promoter sequence of the EBV C promoter (Cp) where it interacts with E2F-1 and Oct-2, and to the family of repeats (FR) region at the latent origin of plasmid replication (oriP) in the EBV plasmid (36). Cp regulates expression of genes required for B cell proliferation in latent EBV infections. The FR regions are upstream of Cp, functioning as an essential enhancer to this promoter, and this interaction is mediated by Epstein-Barr nuclear antigen 1 (EBNA1) (37). Together, these interactions lead to the initiation of transcription of EBV latency proteins (36), suggesting that Bright contributes to maintenance of EBV in certain B cell subsets.

Additional putative gene targets for Bright/ARID3a have been identified as part of the ENCODE project, and that database now lists a number of potential gene targets for Bright/ARID3a in human cell lines (38). One target of particular interest is *Oct4*. We found that Bright binds to the *Oct4* promoter and acts to repress its transcription in mouse embryonic fibroblasts (39). These data are the first to indicate Bright can have repressive as well as activating functions, such as those described for the IgH locus. Further, the *Drosophila* ortholog of Bright, *Drill1*, recruits *Groucho* to *Dorsal* where it is also associated with strong repressive potential (6). It is likely that Bright levels will affect multiple gene targets in hematopoietic cells, where it will be important to consider that it may also act to suppress gene expression.



## CO-REGULATORY AND INTERACTING PROTEINS

Dimerization of Bright was required for binding to the IgH locus in mobility shift assays (8). Earlier antibody supershift assays were the first to suggest that the Bright complex might also contain additional proteins, potentially including topoisomerase II (40). The S107 V<sub>H</sub> gene, containing the prototypic Bright-binding site, was required for immune responses against phosphorylcholine, a response deficient in mice lacking Btk [Ref. (23), reviewed in Ref. (24, 25)]. These data led to experiments demonstrating that Btk interacted directly with Bright, and that its kinase activity was required for Bright-associated transactivation of the IgH promoter (41). Further studies demonstrated that the transcription factor TFII-I, a Btk target [reviewed in Ref. (42)], also bound to Bright through the helix-turn helix domain at the carboxyl end of Bright, and that this interaction was also required for activation of the immunoglobulin locus (33). TFII-I is ubiquitously expressed in multiple isoforms and functions as a transcription factor for basally expressed genes and a transcription activator for upstream protein complexes [reviewed in Ref. (42)]. This suggests that Bright may act as a DNA-binding component to recruit or tether other transcription activating proteins to specific promoter sites.

A truncated form of human Bright/ARID3a was cloned and identified by others from embryonal carcinoma cells as an E2F-binding protein, E2FBP1 (43). These studies showed interactions of Bright and E2F-1, a pRB-controlled protein important for cell cycle regulation, and linked Bright function to cell cycle regulated functions. Functional screens for products that rescued Ras-induced senescence in mouse embryonic fibroblasts also identified Bright (hDrill1) as a candidate protein (44). Those studies also linked Drill1/Bright functions to proliferation and pRB-mediated pathways; however, results using human cells did not mirror those observed in mouse embryonic fibroblasts. In other model systems that required p53, Bright over-expression induced G1 arrest by activating p21 in response to DNA damage in human osteosarcoma cells (45). In this and other studies, ARID3a was found to be a direct downstream target of p53, and a co-regulator with p53 of other gene targets (45, 46). The multiple roles played by Bright/ARID3a during the cell cycle in different cell types highlight the necessity to use caution in interpreting data that may be of a complex nature and the result of interactions with different intracellular mediators. Interpretation of studies involving over-expression of Bright may be further complicated, as levels of Bright within cells of the same lineage appear to be tightly regulated as a consequence of the differentiation state (14). Indeed, SUMOylation of Bright was reported to impair interactions with E2F-1 while promoting transcriptional activation of myeloid lineage-specific genes in HSC populations (47), indicating further complexities that may exist in some cell types due to post-translational modifications of Bright and/or the proteins with which it may interact.

Bright has also been shown to interact with several components of promyelocytic leukemia nuclear bodies (PML NBs), including the ubiquitously expressed protein SP100 and the lymphoid-restricted homolog, LYSp100B (48). Bright was found to colocalize with Sp100 in PML NBs and was found to repress Bright transactivation functions, while LYSp100 strongly stimulated Bright

transactivation of the IgH locus. These data support functions for Bright in higher order chromatin topology and epigenetic regulation. Other studies suggest that levels of Bright/E2FBP1 are important for maintenance of PML NBs and cell viability (49, 50). Some viral proteins, including those from human herpes simplex viruses, also disrupt PML NBs and show linked regulatory effects with Bright/E2FBP1 (51). Clearly, it will be important to understand further consequences of Bright functions in these important nuclear structures.

Finally, as we continue to see examples of transcription factors that interact to form large chromatin modulatory complexes or interactomes that regulate large sets of genes involved in specific cellular processes, it is likely that we will identify new protein partners for Bright. Recent findings suggest Bright/ARID3a is one of a number of genes induced by gamma-interferon in Th1 cells, a T helper cell subset previously unknown to express Bright (52). These findings emphasize again the likelihood that Bright may play distinct regulatory roles in different types of cells.

## REGULATION OF BRIGHT

Id1, a member of a family of three proteins described to be negative regulators of E2A proteins (53, 54), was also shown to interact with Bright/Drill1 in human fetal lung fibroblasts and in lung fibroblasts from patients with idiopathic fibrosis (55). As with other factors inhibited by Id proteins, Id1 formed a complex with Bright that abrogated its DNA-binding activity. In fibrotic lung tissues, Bright was expressed abundantly as a consequence of TGF- $\beta$  signaling (55). Lending further support that Bright may be a downstream mediator of TGF- $\beta$  signaling, the Bright ortholog in *Xenopus* was required for normal development of mesoderm in embryos, through SMAD2-dependent TGF $\beta$  pathways (56). Furthermore, our lab has observed a link between levels of Bright in human B lymphocytes and expression of TGF- $\beta$  pathway associated genes (our unpublished data), additionally supporting associations between TGF- $\beta$  signaling and Bright induction. Id proteins, important for regulation of Bright in lung tissues (55), have also been shown to be important in lineage decisions and in directing B cells toward MZ versus FO B cell phenotypes (57, 58). As mentioned previously, we demonstrated that levels of Bright contribute to those same B cell lineage decisions in transgenic mouse models (27). Therefore, we speculate that Id proteins may also regulate Bright during B cell development.

Very little is known regarding Bright regulation. In B lymphocyte lineage cells, Bright is tightly regulated during differentiation at the level of transcription (14). An important microRNA family regulating transcript levels during hematopoiesis is miRNA125. This microRNA family consists of three members that function at different stages of this process. Myeloid lineage fate decisions have been reported to be regulated by miR125b, pushing granulocyte-macrophage progenitors for the myeloid lineage toward macrophage differentiation (59). Previous studies characterized expression of the miR125 family of micro RNAs in human B lymphocytes at various stages of differentiation, showing members of this family were differentially expressed according to the maturation state of the cells (60). In addition, these studies indicated that over-expression of miR125b inhibited B cell differentiation and affected survival of myeloma cells. Although it is

unclear whether some of these effects may be mediated through suppression of Bright, others have shown that Bright is a direct target of miR125b in B cell progenitors (61). Expression of miR125b in human pre-BI cells increased their proliferation in culture. Similar responses were observed in B cell acute lymphocytic leukemias (B-ALL), where these effects were mediated through suppression of ARID3a (61). Suppression of ARID3a in those cells also resulted in decreased apoptosis in a p53-independent fashion. Interestingly, increased expression of miR125b did not block *in vitro* pre-B differentiation (61). Thus, Bright functions likely differ according to the maturation state of the B cell.

## IMPLICATIONS IN HEALTH AND DISEASE

Because Bright was first identified in B lymphocytes, its functions have been better elucidated in those cells. Over-expression of Bright in mouse B lineage cells increased production of autoantibodies with anti-nuclear antigen (ANA) specificities (26, 27). These antibodies formed immunoglobulin deposits in the kidney glomeruli, although this did not affect kidney function nor did mice display autoimmune phenotypes that threatened mortality. These data imply that Bright over-expression in B lineage cells may predispose those cells toward autoreactive phenotypes, either by expanding B cells with autoimmune phenotypes or by allowing their escape from important tolerance checkpoints. Increased numbers of T1 transitional B cells and MZ B cells have been implicated in autoimmune disease in both mouse and human and both of these B lineage subsets express Bright (26, 27). ANA production is a defining characteristic of autoimmune patients with systemic lupus erythematosus (SLE), and we found that those patients also show increased numbers of Bright<sup>+</sup> peripheral blood B lineage cells that are associated with increased disease activity (our unpublished results). Several studies have linked EBV and SLE, with nearly all pediatric and adult SLE showing EBV infection [reviewed in Ref. (62)]. Intriguingly, Bright is induced in human B cells upon infection with EBV (14), and EBV requires Bright for maintenance of latency genes (36). In addition, many miRNAs have been noted to be differentially expressed in SLE patients versus healthy controls, and miR125a, a member of the family responsible for down-regulation of Bright activity in B cell progenitors (61), was described as being down-regulated in lupus lymphocytes [reviewed in Ref. (63)]. Together, these studies highlight the importance for future mechanistic studies to explore the links between Bright expression and ANA production in SLE patients and their relationship to epigenetic changes implicated in SLE pathology (64).

Bright dysregulation has also been implicated in several types of malignancies, including those derived from hematopoietic lineage cells. Analyses of diffuse large B cell lymphomas (DLBCL) over a decade ago identified two distinct subtypes of those malignancies with different survival advantages, the activated B-like (ABC) and germinal center B-like (GCB) DLBCL by gene expression profiling, which indicated differential ARID3a expression in the two subsets (65). More recent large scale comparative analyses of gene expression patterns in >2000 cases of DLBCL identified Bright/ARID3a as a member of a family of transcription factor signature genes consistently associated with ABC DLBCL, suggesting it might be a useful marker for identification of this subset

of lymphoma (66). To the contrary, in the pre-B cell malignancy, B-ALL, Bright expression was down-regulated compared to levels found in healthy pre-B cells as a consequence of 30- to 600-fold higher expression of miR125b, resulting in the oncogenic properties of increased proliferation and cell survival (61). These data suggest that dysregulated Bright/ARID3a levels may contribute to malignancy. This may not be surprising in light of associations between Bright, p53, and other critical cell cycle mediators described above. In keeping with the multiple advances describing additional cell types expressing ARID3a, recent studies also indicate that high levels of Bright/ARID3a may distinguish colorectal carcinomas with good prognosis and a more differentiated phenotype (67).

Interestingly, a B-ALL patient sample with down-regulated levels of Bright was reported to have significant upregulation of pluripotent factors (61). We previously showed that loss of Bright expression in mouse B lineage cells up-regulated pluripotency gene expression and promoted developmental plasticity, giving rise to cells that resembled induced pluripotent stem cells (10). Many types of malignancies are proposed to contain adult stem cell populations that contribute to their oncogenic potential [reviewed in Ref. (68)]. Stem cells, as a consequence of their self-renewal properties, express many genes commonly dysregulated in oncogenesis. Our recent studies show that Bright knockout mouse embryonic fibroblasts can spontaneously form stem cell-like colonies and the key pluripotency factor, Oct4, is repressed as a consequence of Bright function (39). The ability to reprogram cells from multiple sources, including hematopoietic lineage cells, by manipulating expression of Oct4 and other pluripotency-related transcripts has tremendous potential for regenerative medicine (69). We hypothesize that directed manipulation of Bright levels will also have useful applications in regeneration of some cell types. The next decade is full of Bright promise.

## ACKNOWLEDGMENTS

The authors thank Dr. S. Ferrell, K. Rose, and D. Lamb for data reported in this review. We also thank members of our lab for helpful discussions. In addition, we thank B. Hurt for graphics support and S. Wasson for manuscript preparation. Studies described in this work have been supported by the Oklahoma Center for Adult Stem Cell Research, the Lupus Foundation, and the National Institutes of Health R21AI090343 and R01AI044215 to Carol F. Webb.

## REFERENCES

1. Webb CF, Das C, Eneff KL, Tucker PW. Identification of a matrix-associated region 5' of an immunoglobulin heavy chain variable region gene. *Mol Cell Biol* (1991) 11(10):5206–11.
2. Webb CF, Das C, Eaton S, Calame K, Tucker PW. Novel protein-DNA interactions associated with increased immunoglobulin transcription in response to antigen plus interleukin-5. *Mol Cell Biol* (1991) 11(10):5197–205.
3. Webb CF, Das C, Coffman RL, Tucker PW. Induction of immunoglobulin mu mRNA in a B cell transfectant stimulated with interleukin-5 and a T-dependent antigen. *J Immunol* (1989) 143(12):3934–9.
4. Wilsker D, Probst L, Wain HM, Maltais L, Tucker PW, Moran E. Nomenclature of the ARID family of DNA-binding proteins. *Genomics* (2005) 86(2):242–51. doi:10.1016/j.ygeno.2005.03.013
5. Patsialou A, Wilsker D, Moran E. DNA-binding properties of ARID family proteins. *Nucleic Acids Res* (2005) 33(1):66–80. doi:10.1093/nar/gki145

6. Kortschak RD, Tucker PW, Saint R. ARID proteins come in from the desert. *Trends Biochem Sci* (2000) **25**(6):294–9. doi:10.1016/S0968-0004(00)01597-8
7. Herrscher RF, Kaplan MH, Lelsz DL, Das C, Scheuermann R, Tucker PW. The immunoglobulin heavy-chain matrix-associating regions are bound by Bright: a B cell-specific trans-activator that describes a new DNA-binding protein family. *Genes Dev* (1995) **9**(24):3067–82. doi:10.1101/gad.9.24.3067
8. Nixon JC, Rajaiya J, Webb CF. Mutations in the DNA-binding domain of the transcription factor Bright act as dominant negative proteins and interfere with immunoglobulin transactivation. *J Biol Chem* (2004) **279**(50):52465–72. doi:10.1074/jbc.M403028200
9. Kim D, Probst L, Das C, Tucker PW. REKLES is an ARID3-restricted multifunctional domain. *J Biol Chem* (2007) **282**(21):15768–77. doi:10.1074/jbc.M700397200
10. An G, Miner CA, Nixon JC, Kincade PW, Bryant J, Tucker PW, et al. Loss of Bright/ARID3a function promotes developmental plasticity. *Stem Cells* (2010) **28**(9):1560–7. doi:10.1002/stem.491
11. Wilkinson AC, Gottgens B. Transcriptional regulation of haematopoietic stem cells. *Adv Exp Med Biol* (2013) **786**:187–212. doi:10.1007/978-94-007-6621-1\_11
12. Rosmarin AG, Yang Z, Resendes KK. Transcriptional regulation in myelopoiesis: hematopoietic fate choice, myeloid differentiation, and leukemogenesis. *Exp Hematol* (2005) **33**(2):131–43. doi:10.1016/j.exphem.2004.08.015
13. Ramirez J, Lukin K, Hagman J. From hematopoietic progenitors to B cells: mechanisms of lineage restriction and commitment. *Curr Opin Immunol* (2010) **22**(2):177–84. doi:10.1016/j.coi.2010.02.003
14. Nixon JC, Rajaiya JB, Ayers N, Evetts S, Webb CF. The transcription factor, Bright, is not expressed in all human B lymphocyte subpopulations. *Cell Immunol* (2004) **228**(1):42–53. doi:10.1016/j.cellimm.2004.03.004
15. Webb CF, Smith EA, Medina KL, Buchanan KL, Smithson G, Dou S. Expression of bright at two distinct stages of B lymphocyte development. *J Immunol* (1998) **160**(10):4747–54.
16. Webb CF, Bryant J, Popowski M, Allred L, Kim D, Harriss J, et al. The ARID family transcription factor bright is required for both hematopoietic stem cell and B lineage development. *Mol Cell Biol* (2011) **31**(5):1041–53. doi:10.1128/MCB.01448-10
17. Heng TS, Painter MW. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol* (2008) **9**(10):1091–4. doi:10.1038/nri1008-1091
18. Shinkai Y, Rathbun G, Lam KP, Oltz EM, Stewart V, Mendelsohn M, et al. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* (1992) **68**(5):855–67. doi:10.1016/0092-8674(92)90029-C
19. Spanopoulou E, Roman CA, Corcoran LM, Schlissel MS, Silver DP, Nemazee D, et al. Functional immunoglobulin transgenes guide ordered B-cell differentiation in Rag-1-deficient mice. *Genes Dev* (1994) **8**(9):1030–42. doi:10.1101/gad.8.9.1030
20. Montecino-Rodriguez E, Dorshkind K. B-1 B cell development in the fetus and adult. *Immunity* (2012) **36**(1):13–21. doi:10.1016/j.immuni.2011.11.017
21. Cerutti A, Cols M, Puga I. Marginal zone B cells: virtues of innate-like antibody-producing lymphocytes. *Nat Rev Immunol* (2013) **13**(2):118–32. doi:10.1038/nri3383
22. Nixon JC, Ferrell S, Miner C, Oldham AL, Hochgeschwender U, Webb CF. Transgenic mice expressing dominant-negative bright exhibit defects in B1 B cells. *J Immunol* (2008) **181**(10):6913–22.
23. Brown M, Stenzel-Poore M, Rittenberg MB. Immunologic memory to phosphocholine. VII. Lack of T15 V1 gene utilization in Xid anti-PC hybridomas. *J Immunol* (1985) **135**(5):3558–63.
24. Mohamed AJ, Yu L, Bäckesjö CM, Vargas L, Faryal R, Aints A, et al. Bruton's tyrosine kinase (Btk): function, regulation, and transformation with special emphasis on the PH domain. *Immunol Rev* (2009) **228**(1):58–73. doi:10.1111/j.1600-065X.2008.00741.x
25. Khan WN, Alt FW, Gerstein RM, Malynn BA, Larsson I, Rathbun G, et al. Defective B cell development and function in Btk-deficient mice. *Immunity* (1995) **3**(3):283–99. doi:10.1016/1074-7613(95)90114-0
26. Shankar M, Nixon JC, Maier S, Workman J, Farris AD, Webb CF. Anti-nuclear antibody production and autoimmunity in transgenic mice that overexpress the transcription factor Bright. *J Immunol* (2007) **178**(5):2996–3006.
27. Oldham AL, Miner CA, Wang HC, Webb CF. The transcription factor Bright plays a role in marginal zone B lymphocyte development and autoantibody production. *Mol Immunol* (2011) **49**(1–2):367–79. doi:10.1016/j.molimm.2011.09.008
28. Kin NW, Crawford DM, Liu J, Behrens TW, Kearney JE. DNA microarray gene expression profile of marginal zone versus follicular B cells and idiotype positive marginal zone B cells before and after immunization with *Streptococcus pneumoniae*. *J Immunol* (2008) **180**(10):6663–74.
29. Schmidt C, Kim D, Ippolito GC, Naqvi HR, Probst L, Mathur S, et al. Signalling of the BCR is regulated by a lipid rafts-localised transcription factor, Bright. *EMBO J* (2009) **28**(6):711–24. doi:10.1038/emboj.2009.20
30. Johnston CM, Wood AL, Bolland DJ, Corcoran AE. Complete sequence assembly and characterization of the C57BL/6 mouse Ig heavy chain V region. *J Immunol* (2006) **176**(7):4221–34.
31. Corcoran A, Crowley B, Dewhurst C, Pizer BL, Doyle S. Establishment of functional B cell memory against parvovirus B19 capsid proteins may be associated with resolution of persistent infection. *J Med Virol* (2006) **78**(1):125–8. doi:10.1002/jmv.20513
32. Goebel P, Montalbano A, Ayers N, Kompfner E, Dickinson L, Webb CF, et al. High frequency of matrix attachment regions and cut-like protein x/CCAAT-displacement protein and B cell regulator of IgH transcription binding sites flanking Ig V region genes. *J Immunol* (2002) **169**(5):2477–87.
33. Rajaiya J, Nixon JC, Ayers N, Desgranges ZP, Roy AL, Webb CF. Induction of immunoglobulin heavy-chain transcription through the transcription factor Bright requires TFII-I. *Mol Cell Biol* (2006) **26**(12):4758–68. doi:10.1128/MCB.02009-05
34. Lin D, Ippolito GC, Zong RT, Bryant J, Koslovsky J, Tucker P. Bright/ARID3A contributes to chromatin accessibility of the immunoglobulin heavy chain enhancer. *Mol Cancer* (2007) **6**:23. doi:10.1186/1476-4598-6-23
35. Kaplan MH, Zong RT, Herrscher RF, Scheuermann RH, Tucker PW. Transcriptional activation by a matrix associating region-binding protein: contextual requirements for the function of bright. *J Biol Chem* (2001) **276**(24):21325–30. doi:10.1074/jbc.M100836200
36. Boreström C, Forsman A, Rüetschi U, Rymo L. E2F1, ARID3A/Bright and Oct-2 factors bind to the Epstein-Barr virus C promoter, EBNA1 and oriP, participating in long-distance promoter-enhancer interactions. *J Gen Virol* (2012) **93**(Pt 5):1065–75. doi:10.1099/vir.0.038752-0
37. Frappier L, O'Donnell M. Epstein-Barr nuclear antigen 1 mediates a DNA loop within the latent replication origin of Epstein-Barr virus. *Proc Natl Acad Sci U S A* (1991) **88**(23):10875–9. doi:10.1073/pnas.88.23.10875
38. Karolchik D, Barber GP, Casper J, Clawson H, Cline MS, Diekhans M, et al. The UCSC Genome Browser database: 2014 update. *Nucleic Acids Res* (2014) **42**(1):D764–70. doi:10.1093/nar/gkt1168
39. Popowski M, Templeton TD, Lee BK, Rhee C, Li H, Miner C, et al. Bright/Arld3A acts as a barrier to somatic cell reprogramming through direct regulation of Oct4, Sox2, and Nanog. *Stem Cell Reports* (2014) **2**(1):26–35. doi:10.1016/j.stemcr.2013.12.002
40. Webb CF, Eneff KL, Drake FH. A topoisomerase II-like protein is part of an inducible DNA-binding protein complex that binds 5' of an immunoglobulin promoter. *Nucleic Acids Res* (1993) **21**(18):4363–8. doi:10.1093/nar/21.18.4363
41. Rajaiya J, Hatfield M, Nixon JC, Rawlings DJ, Webb CF. Bruton's tyrosine kinase regulates immunoglobulin promoter activation in association with the transcription factor Bright. *Mol Cell Biol* (2005) **25**(6):2073–84. doi:10.1128/MCB.25.6.2073-2084.2005
42. Roy AL. Biochemistry and biology of the inducible multifunctional transcription factor TFII-I. *Gene* (2001) **274**(1–2):1–13. doi:10.1016/S0378-1119(01)00625-4
43. Suzuki M, Okuyama S, Okamoto S, Shirasuna K, Nakajima T, Hachiya T, et al. A novel E2F binding protein with Myc-type HLH motif stimulates E2F-dependent transcription by forming a heterodimer. *Oncogene* (1998) **17**(7):853–65. doi:10.1038/sj.onc.1202163
44. Peeper DS, Shvarts A, Brummelkamp T, Douma S, Koh EY, Daley GQ, et al. A functional screen identifies hDRIL1 as an oncogene that rescues RAS-induced senescence. *Nat Cell Biol* (2002) **4**(2):148–53. doi:10.1038/ncb742
45. Lestari W, Ichwan SJ, Otsu M, Yamada S, Iseki S, Shimizu S, et al. Cooperation between ARID3A and p53 in the transcriptional activation of p21WAF1 in

- response to DNA damage. *Biochem Biophys Res Commun* (2012) **417**(2):710–6. doi:10.1016/j.bbrc.2011.12.003
46. Ma K, Araki K, Ichwan SJ, Suganuma T, Tamamori-Adachi M, Ikeda MA. E2FBP1/DRIL1, an AT-rich interaction domain-family transcription factor, is regulated by p53. *Mol Cancer Res* (2003) **1**(6):438–44.
  47. Prieur A, Nacerddine K, van Lohuizen M, Peeper DS. SUMOylation of DRIL1 directs its transcriptional activity towards leukocyte lineage-specific genes. *PLoS One* (2009) **4**(5):e5542. doi:10.1371/journal.pone.0005542
  48. Zong RT, Das C, Tucker PW. Regulation of matrix attachment region-dependent, lymphocyte-restricted transcription through differential localization within promyelocytic leukemia nuclear bodies. *EMBO J* (2000) **19**(15):4123–33. doi:10.1093/emboj/19.15.4123
  49. Fukuyo Y, Takahashi A, Hara E, Horikoshi N, Pandita TK, Nakajima T. E2FBP1 antagonizes the p16(INK4A)-Rb tumor suppressor machinery for growth suppression and cellular senescence by regulating promyelocytic leukemia protein stability. *Int J Oral Sci* (2011) **3**(4):200–8. doi:10.4248/IJOS11071
  50. Fukuyo Y, Mogi K, Tsunematsu Y, Nakajima T. E2FBP1/hDril1 modulates cell growth through downregulation of promyelocytic leukemia bodies. *Cell Death Differ* (2004) **11**(7):747–59. doi:10.1038/sj.cdd.4401412
  51. Fukuyo Y, Horikoshi N, Ishov AM, Silverstein SJ, Nakajima T. The herpes simplex virus immediate-early ubiquitin ligase ICP0 induces degradation of the ICP0 repressor protein E2FBP1. *J Virol* (2011) **85**(7):3356–66. doi:10.1128/JVI.02105-10
  52. Hakim O, Sung MH, Nakayama S, Voss TC, Baek S, Hager GL. Spatial congregation of STAT binding directs selective nuclear architecture during T-cell functional differentiation. *Genome Res* (2013) **23**(3):462–72. doi:10.1101/gr.147652.112
  53. Loveys DA, Streiff MB, Kato GJ. E2A basic-helix-loop-helix transcription factors are negatively regulated by serum growth factors and by the Id3 protein. *Nucleic Acids Res* (1996) **24**(14):2813–20. doi:10.1093/nar/24.14.2813
  54. Sun XH, Copeland NG, Jenkins NA, Baltimore D. Id proteins Id1 and Id2 selectively inhibit DNA binding by one class of helix-loop-helix proteins. *Mol Cell Biol* (1991) **11**(11):5603–11.
  55. Lin L, Zhou Z, Zheng L, Alber S, Watkins S, Ray P, et al. Cross talk between Id1 and its interactive protein Dril1 mediate fibroblast responses to transforming growth factor-beta in pulmonary fibrosis. *Am J Pathol* (2008) **173**(2):337–46. doi:10.2353/ajpath.2008.070915
  56. Callery EM, Smith JC, Thomsen GH. The ARID domain protein dril1 is necessary for TGF(beta) signaling in *Xenopus* embryos. *Dev Biol* (2005) **278**(2):542–59. doi:10.1016/j.ydbio.2004.11.017
  57. Zhang P, Zhao Y, Sun XH. Notch-regulated periphery B cell differentiation involves suppression of E protein function. *J Immunol* (2013) **191**(2):726–36. doi:10.4049/jimmunol.1202134
  58. Quong MW, Martensson A, Langerak AW, Rivera RR, Nemazee D, Murre C. Receptor editing and marginal zone B cell development are regulated by the helix-loop-helix protein, E2A. *J Exp Med* (2004) **199**(8):1101–12. doi:10.1084/jem.20031180
  59. Shaham L, Binder V, Gefen N, Borkhardt A, Izraeli S. MiR-125 in normal and malignant hematopoiesis. *Leukemia* (2012) **26**(9):2011–8. doi:10.1038/leu.2012.90
  60. Gururajan M, Haga CL, Das S, Leu CM, Hodson D, Josson S, et al. MicroRNA 125b inhibition of B cell differentiation in germinal centers. *Int Immunol* (2010) **22**(7):583–92. doi:10.1093/intimm/dxq042
  61. Puissegur MP, Eichner R, Quelen C, Coyaude E, Mari B, Lebrigand K, et al. B-cell regulator of immunoglobulin heavy-chain transcription (Bright)/ARID3a is a direct target of the oncomir microRNA-125b in progenitor B-cells. *Leukemia* (2012) **26**(10):2224–32. doi:10.1038/leu.2012.95
  62. Draborg AH, Duus K, Houen G. Epstein-Barr virus and systemic lupus erythematosus. *Clin Dev Immunol* (2012) **2012**:370516. doi:10.1155/2012/370516
  63. Amarilio G, La Cava A. miRNA in systemic lupus erythematosus. *Clin Immunol* (2012) **144**(1):26–31. doi:10.1016/j.clim.2012.04.005
  64. Frangou EA, Bertsias GK, Boumpas DT. Gene expression and regulation in systemic lupus erythematosus. *Eur J Clin Invest* (2013) **43**(10):1084–96. doi:10.1111/eci.12130
  65. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* (2000) **403**(6769):503–11. doi:10.1038/35000501
  66. Care MA, Barrans S, Worrillow L, Jack A, Westhead DR, Tooze RM. A microarray platform-independent classification tool for cell of origin class allows comparative analysis of gene expression in diffuse large B-cell lymphoma. *PLoS One* (2013) **8**(2):e55895. doi:10.1371/journal.pone.0055895
  67. Song M, Kim H, Kim WK, Hong SP, Lee C, Kim H. High expression of AT-rich interactive domain 3A (ARID3A) is associated with good prognosis in colorectal carcinoma. *Ann Surg Oncol* (2013). doi:10.1245/s10434-013-3435-2
  68. Bao B, Ahmad A, Azmi AS, Ali S, Sarkar FH. Overview of cancer stem cells (CSCs) and mechanisms of their regulation: implications for cancer therapy. *Curr Protoc Pharmacol* (2013) **Chapter 14**:Unit1425. doi:10.1002/0471141755.ph1425s61
  69. Tavernier G, Mlody B, Demeester J, Adjaye J, De Smedt SC. Current methods for inducing pluripotency in somatic cells. *Adv Mater* (2013) **25**(20):2765–71. doi:10.1002/adma.201204874

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 January 2014; paper pending published: 11 February 2014; accepted: 04 March 2014; published online: 19 March 2014.

Citation: Ratliff ML, Templeton TD, Ward JM and Webb CF (2014) The Bright side of hematopoiesis: regulatory roles of ARID3a/Bright in human and mouse hematopoiesis. *Front. Immunol.* 5:113. doi: 10.3389/fimmu.2014.00113

This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Ratliff, Templeton, Ward and Webb. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Architecture and expression of the *Nfatc1* gene in lymphocytes

Ronald Rudolf<sup>1†</sup>, Rhoda Busch<sup>1†</sup>, Amiya K. Patra<sup>1</sup>, Khalid Muhammad<sup>1</sup>, Andris Avots<sup>1</sup>, Jean-Christophe Andrau<sup>2</sup>, Stefan Klein-Hessling<sup>1</sup> and Edgar Serfling<sup>1\*</sup>

<sup>1</sup> Department of Molecular Pathology, Comprehensive Cancer Center Mainfranken, Institute of Pathology, University of Würzburg, Würzburg, Germany

<sup>2</sup> Centre d'Immunologie de Marseille-Luminy, Université Aix-Marseille, Marseille, France

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

John D. Colgan, University of Iowa, USA

James Hagman, National Jewish Health, USA

## \*Correspondence:

Edgar Serfling, Department of Molecular Pathology, Institute of Pathology, University of Würzburg, Josef-Schneider-Str. 2, Würzburg D-97080, Germany  
e-mail: serfling.e@mail.uni-wuerzburg.de

<sup>†</sup> Ronald Rudolf and Rhoda Busch have contributed equally to this work.

In lymphocytes, the three NFAT factors NFATc1 (also designated as NFAT2), NFATc2 (NFAT1), and NFATc3 (NFAT4 or NFATx) are expressed and are the targets of immune receptor signals, which lead to a rapid rise of intracellular  $\text{Ca}^{++}$ , the activation of phosphatase calcineurin, and to the activation of cytosolic NFATc proteins. In addition to rapid activation of NFAT factors, immune receptor signals lead to accumulation of the short NFATc1/ $\alpha$ A isoform in lymphocytes which controls their proliferation and survival. In this mini-review, we summarize our current knowledge on the structure and transcription of the *Nfatc1* gene in lymphocytes, which is controlled by two promoters, two poly A addition sites and a remote downstream enhancer. The *Nfatc1* gene resembles numerous primary response genes (PRGs) induced by LPS in macrophages. Similar to the PRG promoters, the *Nfatc1* promoter region is organized in CpG islands, forms DNase I hypersensitive sites, and is marked by histone tail modifications before induction. By studying gene induction in lymphocytes in detail, it will be important to elucidate whether the properties of the *Nfatc1* induction are not only typical for the *Nfatc1* gene but also for other transcription factor genes expressed in lymphocytes.

**Keywords: chromatin, induction, lymphocytes, *Nfatc1*, transcription**

## INTRODUCTION

In peripheral B lymphocytes, the NFATc factors NFATc1, c2, and c3 are the final targets of B cell receptor (BCR)-mediated activation, and inhibiting their induction by the immunosuppressant Cyclosporin A (CsA) abrogates the antigen-induced proliferation of B cells (1). In freshly isolated (naive) splenic B cells according to the number of RNA reads in RNA Seq assays, 10-fold more transcripts were detected for the *Nfatc1* and *Nfatc3* genes than for the *Nfatc2* gene. BCR signals increase the transcription of the *Nfatc1* gene, but not of the *Nfatc2* and *Nfatc3* genes (Muhammad et al., submitted). Although all three NFATc transcription factors (TFs) bind to similar DNA motifs and transactivate the promoters of numerous genes in transfection studies, inactivating the individual *Nfatc* genes in mice resulted in quite diverging phenotypes. Whereas inactivation of the *Nfatc1* gene led to an early death of mice embryos (2, 3), *Nfatc2*<sup>-/-</sup> mice were born at normal Mendelian ratio but developed with age, a hyper-proliferative syndrome and elevated immune responses (4–6). These features of the *Nfatc2*<sup>-/-</sup> mice were found to be accelerated in mice deficient in both NFATc2 and NFATc3 (7). Ablation of NFATc1 in B cells led to a marked reduction in BCR-mediated proliferation and  $\text{Ca}^{++}$  flux, increase in activation induced cell death (AICD), and defects in antibody production upon immunization, whereas opposite effects were observed for *Nfatc2*<sup>-/-</sup> B cells (1, 8).

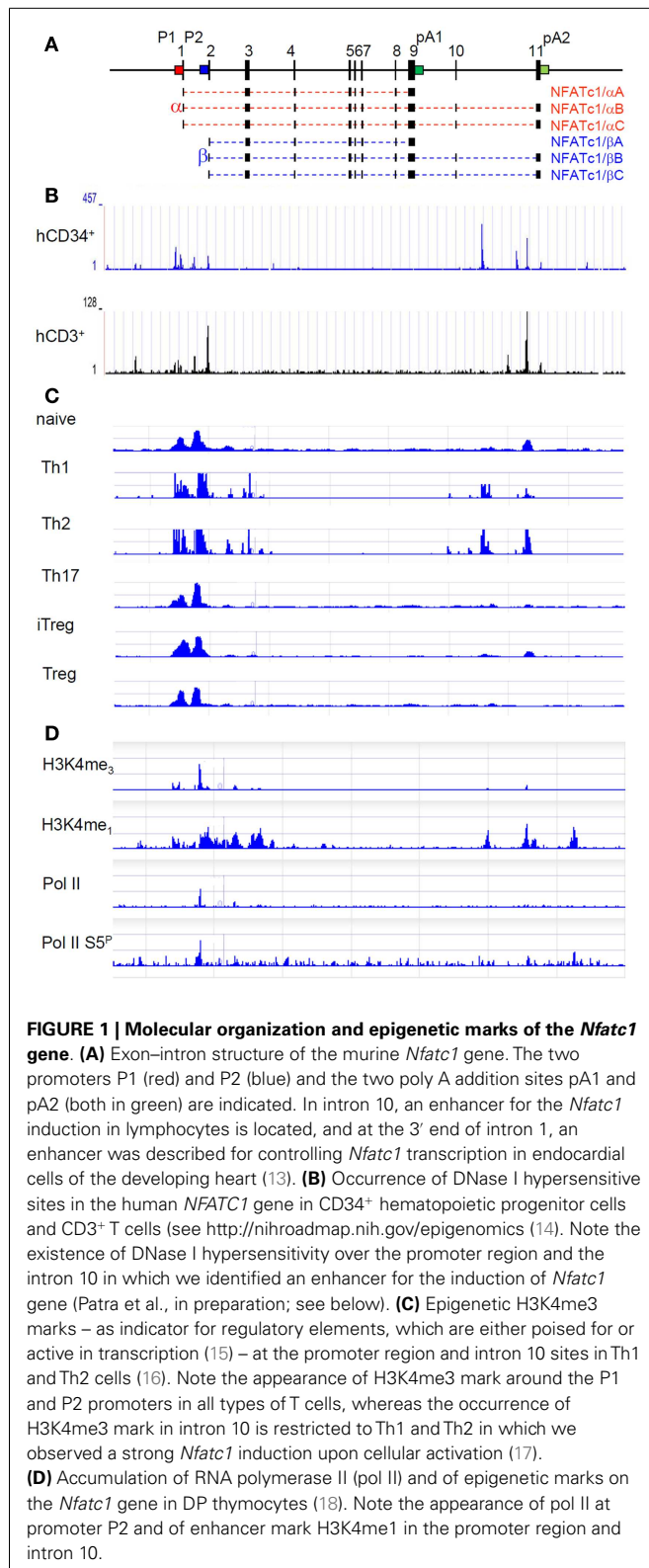
These functional differences between NFATc1 and NFATc2 might be due to the synthesis of NFATc1/ $\alpha$ A, a short isoform of NFATc1, which lacks the C-terminal peptide of approximately 250 amino acids residues typical for most of the other NFATc proteins.

NFATc1/ $\alpha$ A is the most prominent NFAT protein in effector B cells and is able to rescue B cells from early cell death (1).

## STRUCTURE OF THE *Nfatc1* GENE

The genes encoding NFATc1 in mouse and man consist of 11 exons and span approximately 110 and 134 kb DNA, respectively. Due to the existence of two promoters, two poly A sites and alternate splicing events, six NFATc1 RNAs, and proteins are generated in peripheral lymphocytes (9–11) (Figure 1A). The two *Nfatc1* promoters, P1 and P2, show the typical features of eukaryotic promoters. They are highly conserved between mouse and man over 800 bp (P1) or 100 bp (P2) DNA and form DNase I hypersensitive chromatin sites. Both promoters are organized in CpG islands. While in peripheral blood lymphocytes, in Jurkat T cells, and in other lymphoid cell lines in which NFATc1 is expressed the DNA of promoter islands is de-methylated, inactivation of human *NFATC1* gene in several Hodgkin's lymphoma cells lines is correlated with the methylation of all CpG dinucleotides within the P1 promoter (12).

The inducible P1 promoter of 800 bp can be divided into two DNA “homology blocks” (10) of approximately 250 bp DNA, which harbor DNA binding motifs for Sp1 at their termini. Sp1 binding is known to protect CpG islands from DNA methylation (19, 20), and the relatively weak binding of Sp1 to these P1 sites in Hodgkin's lymphoma cells in which the promoter is suppressed led us to speculate that they could function as “road blocks” to prevent the methylation of P1 DNA in effector lymphocytes (12) [but see also Ref. (21) and below for the function of



Sp1 in the control of primary response promoters]. The TF binding motifs within each block of homology represent composite sites for the inducible binding of Creb/Fos/ATF and NF- $\kappa$ B/NFAT

factors. When used as probes in electrophoretic band shift assays (EMSA), the NF- $\kappa$ B/NFAT sites form predominantly NF- $\kappa$ B complexes with proteins from activated T and B cells. However, under conditions of high nuclear NFAT levels these sites can also be bound by NFAT. Together with the two NFAT sites within the distal block of homology, which are arranged in tandem, they enable strong binding of NFAT factors to the P1 promoter (9). This contributes to the NFATc1-mediated auto-regulation of P1-directed *Nfatc1* transcription, which keeps high levels of NFATc1 for days under persistent immune receptor stimulation (11).

In spite of the tight assembly of TF binding sites within the P1 promoter, when transfected into EL-4 thymoma cells, P1-directed luciferase reporter constructs showed a poor induction, which differs markedly from the induction of endogenous *Nfatc1* gene (9). Instead of being induced as the endogenous *Nfatc1* gene by the phorbol ester TPA and the Ca<sup>++</sup>-ionophore ionomycin (which mimic immune receptor signals), inducers of protein kinase A, such as forskolin, led in combination with ionomycin to the strongest induction of P1 in EL-4 thymoma cells (9).

These functional studies on the induction of *Nfatc1* P1 promoter led us to conclude that not all sequence elements controlling the induction of the *Nfatc1* gene in lymphocytes are part of its promoter region. Whereas fusion of more upstream DNA to the highly conserved P1 region of approximately 800 bp did not result in any increase in promoter induction, fusion of a 1-kb DNA fragment from the central region of intron 1 of the *Nfatc1* gene to P1 enhanced its overall activity fourfold to fivefold, but did not affect its induction mode (9). However, when we inserted an element from intron 10 of the *Nfatc1* gene into a luciferase construct directed by the P1 (or P2) promoter, we observed both a strong increase in promoter induction and a mode of induction similar to the endogenous *Nfatc1* gene. These findings suggest that an enhancer for the optimal induction of *Nfatc1* gene is located in intron 10, which supports *Nfatc1* induction in lymphocytes (Patra et al., in preparation).

In Figures 1B,C, mapping studies of DNase I hypersensitive sites in human CD34<sup>+</sup> lymphoid progenitor cells and CD3<sup>+</sup> T cells and of H3K4me3 mark in various subsets of murine T cells are presented for the *Nfatc1* gene. They show that, in addition to the promoter region, both DNase I hypersensitive chromatin sites and H3K4me3 marks were mapped within intron 10 of *Nfatc1*. The distal intron 10 site was found to be marked by H3K4me1 and H3K4me3 modifications (see Figures 1C,D), and the enrichment of H3K4me3 was identified as a feature of active enhancers in T cells (22). And indeed, we determined this site (designated as E2) as an enhancer element that supports the induction of P1 and P2 promoters in lymphocytes (Patra et al., in preparation). Interestingly, E2 appears to be less active (or inactive) in Th17 cells, in thymus-derived regulatory T cells (Treg), and in induced Treg (iTreg) in which we observed a weak *Nfatc1* induction (17).

In resting CD4<sup>+</sup> T cells and DP thymocytes in which *Nfatc1* is poorly expressed, the (P2) promoter region of the *Nfatc1* gene shows characteristics of a “transcription initiation platform.” In ChIP-Seq assays using DP thymocytes, the RNA polymerase II (pol II) was found to be bound at P2 (and not at P1), whereas the enhancer mark H3K4me1 was detected over the entire promoter region and several intron 10 sites (Figure 1D) (18). However, when

ChIP assays were performed for H3K27ac, a mark for active – and not only poised – enhancers (23), only a peak over the central intron 10 enhancer segment E2 appeared in double-negative (DN) thymocytes (Andrau, in preparation) in which the *Nfatc1* gene is expressed more robustly than in DP thymocytes (17).

### Nfatc1 EXPRESSION IN PERIPHERAL B CELLS

When splenic B cells are induced by  $\alpha$ -IgM for 24 h *ex vivo*, the predominant synthesis of short NFATc1 isoform NFATc1/ $\alpha$ A is observed (1, 17). While in Western blots using whole B cell protein a strong, more than 50-fold induction of NFATc1/ $\alpha$ A protein is detected, in real time PCR assays measuring the levels of *Nfatc1*/ $\alpha$ A RNA a 5- to 10-fold increase was observed, and in recent RNA Seq assays, high levels of NFATc1 RNA were found in non-stimulated primary splenic B cells, which are not reflected at the protein level. These observations suggest the existence of both transcriptional and post-transcriptional control mechanisms, which shape the appearance of NFATc1 protein(s) upon B (and T) cell induction.

To study the expression of *Nfatc1* gene at the transcriptional level *in vivo*, we generated a BAC transgenic (tg) mouse line, which expresses an *Egfp* reporter gene under the control of the entire *Nfatc1* locus (17). Within the BAC construct, the *Egfp* reporter replaces exon 3 of the *Nfatc1* gene followed by a SV40 poly A addition signal, which gives rise to short chimeric *Nfatc1/Egfp* RNAs and proteins. Therefore, the *Nfatc1/Egfp* transcripts are generated under the control of all regulatory elements of the *Nfatc1* gene, including both promoters and the downstream enhancer. However, the post-transcriptional mechanisms leading to NFATc1/ $\alpha$ A protein differ certainly between “normal” *Nfatc1* and *Nfatc1/Egfp* transcripts. Thus, in lymphocytes of tg *Nfatc1/Egfp* mice, the expression of chimeric *Nfatc1/Egfp* tg should reflect the transcription of *Nfatc1* locus, but not the expression of NFATc1 proteins.

In tg *Nfatc1/Egfp* mice, the *Nfatc1* gene is expressed as early as in DN thymocytes and in naïve resting T and B cells of peripheral lymphoid organs. Although before the induction of pre-T cell receptor at the transition of DN3 to DN4 thymocytes, NFATc1  $\alpha$ -isoforms are not generated and, therefore, the P1 promoter is less active (or inactive), the *Nfatc1* gene appears to be transcribed at a relatively high level in DN thymocytes lacking any immune receptor (Patra et al., in preparation). This appears also to be the case in naïve and resting T and B lymphocytes. Thus, similar to other TF genes encoding Fos, Jun, Egr, ATF, and further TF factors, which harbor CpG islands in their promoters, the *Nfatc1* gene seems to belong to the group of primary response genes (PRGs) that show a moderate 5- to 10-fold induction upon cellular stimulation. Contrary to secondary response genes (SRGs), which are often induced more than 100-fold, PRGs appear to be organized in an “open” chromatin, which is poised for transcription or transcribed at a low level (24, 25).

To a large part, our current view on the regulation of inducible genes bases on studies about LPS-mediated gene induction in macrophages (21, 26), see also (27) and (24). Previous approaches on the knock down of components of SWI/SNF nucleosome remodeling complexes in macrophages showed that in contrast to the SWI/SNF-dependent induction of SRGs, the LPS-mediated

induction of PRGs is independent of SWI/SNF (28). Similar to promoters of many house-keeping genes, which are also organized in CpG islands, PRG promoters exhibit constitutively active chromatin with unstable nucleosomes, which form constitutive DNase I hypersensitive regions (26). Before induction, they are associated with the initiating version of pol II phosphorylated at “Ser5” within their C-terminal domain (CTD), and with Sp1, which helps to recruit pol II. But contrary to the heat shock genes in *Drosophila*, which are also pre-loaded with pol II and transcribed into short RNAs (29), PRG transcripts in macrophages are elongated to full-length transcripts, which appear to be instable and un-spliced (21). LPS stimulation, however, which often leads to binding of NF- $\kappa$ B to the promoters of PRGs results in the phosphorylation of pol II at position S2 within its CTD repeats and the generation of stable RNAs, which are spliced and processed (21, 26).

The architecture of the *Nfatc1* promoter region and its induction is similar, but not in all aspects, to PRG promoters and their induction in macrophages. In lymphocytes, induction of the *Nfatc1* gene is controlled predominantly by immune receptor signals but not by LPS [or other co-stimulatory signals; see Ref. (1)]. The *Nfatc1* promoter region is organized in CpG islands, forms DNase I hypersensitive sites, and is bound by Sp1 [and CREB, which controls activity-dependent PRG regulation in neurons (30)] prior to its induction by NF- $\kappa$ B. However, induction of the *Nfatc1* promoter differs significantly from that of PRG promoters in macrophages. In contrast to PRGs (31) and similar to the “inducible house-keeping” *Nfkbia* gene (21), the *Nfatc1* gene is efficiently transcribed in lymphocytes prior to the appearance of stable, spliced transcripts in response to receptor signals. These transcripts, however, remain un-translated (Muhammad et al., in preparation.).

### SUMMARY AND IMPLICATIONS

The immune receptor-mediated induction of NFATc1 TFs in peripheral lymphocytes can be divided in two events: (i) the rapid nuclear transport and activation of pre-formed cytosolic NFATc proteins, and (ii) the massive transcriptional and post-transcriptional induction of NFATc1/ $\alpha$ A, a short NFATc1 protein, which differs in many properties from other NFATc proteins (10). Although the induction of the *Nfatc1* gene leading to NFATc1/ $\alpha$ A in lymphocytes resembles the LPS-mediated induction of PRGs in macrophages, it appears to differ from the induction of many PRGs by (i) its high constitutive transcription into spliced transcripts and (ii) its enhancer-mediated control. While the molecular details of these events remain to be elucidated, it will be important to investigate whether the properties of *Nfatc1* induction are specific for the *Nfatc1* gene or a property of immune receptor-mediated induction of many TF genes in lymphocytes. In any way, the detailed knowledge of molecular mechanisms controlling the induction of NFATc1 in lymphocytes could pave the way to interfere with its induction, which controls numerous aspects of adaptive immunity.

### ACKNOWLEDGMENT

This publication was funded by the German Research Foundation (DFG) and the University of Wuerzburg in the funding programme Open Access Publishing.

## REFERENCES

- Bhattacharyya S, Deb J, Patra AK, Thuy Pham DA, Chen W, Vaeth M, et al. NFATc1 affects mouse splenic B cell function by controlling the calcineurin – NFAT signaling network. *J Exp Med* (2011) **208**:823–39. doi:10.1084/jem.20100945
- de la Pompa JL, Timmerman LA, Takimoto H, Yoshida H, Elia AJ, Samper E, et al. Role of the NF-ATc transcription factor in morphogenesis of cardiac valves and septum. *Nature* (1998) **392**:182–6. doi:10.1038/32419
- Ranger AM, Grusby MJ, Hodge MR, Gravalles EM, de la Brousse FC, Hoey T, et al. The transcription factor NF-ATc is essential for cardiac valve formation. *Nature* (1998) **392**:186–90. doi:10.1038/32426
- Hodge MR, Ranger AM, Charles de la Brousse F, Hoey T, Grusby MJ, Glimcher LH. Hyperproliferation and dysregulation of IL-4 expression in NF-ATp-deficient mice. *Immunity* (1996) **4**:397–405. doi:10.1016/S1074-7613(00)80253-8
- Schuh K, Kneitz B, Heyer J, Bommhardt U, Jankevics E, Berberich-Siebelt F, et al. Retarded thymic involution and massive germinal center formation in NF-ATp-deficient mice. *Eur J Immunol* (1998) **28**:2456–66. doi:10.1002/(SICI)1521-4141(199808)28:08<2456::AID-IMMU2456>3.0.CO;2-9
- Xanthoudakis S, Viola JP, Shaw KT, Luo C, Wallace JD, Bozza PT, et al. An enhanced immune response in mice lacking the transcription factor NFAT1. *Science* (1996) **272**:892–5. doi:10.1126/science.272.5263.892
- Ranger AM, Oukka M, Rengarajan J, Glimcher LH. Inhibitory function of two NFAT family members in lymphoid homeostasis and Th2 development. *Immunity* (1998) **9**:627–35. doi:10.1016/S1074-7613(00)80660-3
- Samanta DN, Palmethofer A, Marinkovic D, Wirth T, Serfling E, Nitschke L. B cell hyperresponsiveness and expansion of mature follicular B cells but not of marginal zone B cells in NFATc2/c3 double-deficient mice. *J Immunol* (2005) **174**:4797–802.
- Chuvpilo S, Jankevics E, Tyrsin D, Akimzhanov A, Moroz D, Jha MK, et al. Autoregulation of NFATc1/A expression facilitates effector T cells to escape from rapid apoptosis. *Immunity* (2002) **16**:881–95. doi:10.1016/S1074-7613(02)00329-1
- Serfling E, Avots A, Klein-Hessling S, Rudolf R, Vaeth M, Berberich-Siebelt F. NFATc1/alphaA: the other face of NFAT factors in lymphocytes. *Cell Commun Signal* (2012) **10**:16. doi:10.1186/1478-811X-10-16
- Serfling E, Chuvpilo S, Liu J, Hofer T, Palmethofer A. NFATc1 autoregulation: a crucial step for cell-fate determination. *Trends Immunol* (2006) **27**:461–9. doi:10.1016/j.it.2006.08.005
- Akimzhanov A, Krenacs L, Schlegel T, Klein-Hessling S, Bagdi E, Stelkovics E, et al. Epigenetic changes and suppression of the nuclear factor of activated T cell 1 (NFATC1) promoter in human lymphomas with defects in immunoreceptor signaling. *Am J Pathol* (2008) **172**:215–24. doi:10.2353/ajpath.2008.070294
- Zhou B, Wu B, Tompkins KL, Boyer KL, Grindley JC, Baldwin HS. Characterization of Nfatc1 regulation identifies an enhancer required for gene expression that is specific to pro-valve endocardial cells in the developing heart. *Development* (2005) **132**:1137–46. doi:10.1242/dev.01640
- Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, et al. The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* (2010) **28**:1045–8. doi:10.1038/nbt1010-1045
- Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, et al. Active genes are tri-methylated at K4 of histone H3. *Nature* (2002) **419**:407–11. doi:10.1038/nature01080
- Wei G, Wei L, Zhu J, Zang C, Hu-Li J, Yao Z, et al. Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4+ T cells. *Immunity* (2009) **30**:155–67. doi:10.1016/j.immuni.2008.12.009
- Hock M, Vaeth M, Rudolf R, Patra AK, Pham DA, Muhammad K, et al. NFATc1 induction in peripheral T and B lymphocytes. *J Immunol* (2013) **190**:2345–54. doi:10.4049/jimmunol.1201591
- Koch F, Fenouil R, Gut M, Cauchy P, Albert TK, Zacarias-Cabeza J, et al. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat Struct Mol Biol* (2011) **18**:956–63. doi:10.1038/nsmb.2085
- Brandeis M, Frank D, Keshet I, Siegfried Z, Mendelsohn M, Nemes A, et al. Sp1 elements protect a CpG island from de novo methylation. *Nature* (1994) **371**:435–8. doi:10.1038/371435a0
- Macleod D, Charlton J, Mullins J, Bird AP. Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. *Genes Dev* (1994) **8**:2282–92. doi:10.1101/gad.8.19.2282
- Hargreaves DC, Horng T, Medzhitov R. Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* (2009) **138**:129–45. doi:10.1016/j.cell.2009.05.047
- Pekowska A, Benoukraf T, Zacarias-Cabeza J, Belhocine M, Koch F, Holota H, et al. H3K4 tri-methylation provides an epigenetic signature of active enhancers. *EMBO J* (2011) **30**:4198–210. doi:10.1038/emboj.2011.295
- Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* (2010) **107**:21931–6. doi:10.1073/pnas.1016071107
- Fowler T, Sen R, Roy AL. Regulation of primary response genes. *Mol Cell* (2011) **44**:348–60. doi:10.1016/j.molcel.2011.09.014
- Smale ST. Transcriptional regulation in the innate immune system. *Curr Opin Immunol* (2012) **24**:51–7. doi:10.1016/j.coi.2011.12.008
- Ramirez-Carrozzi VR, Braas D, Bhatt DM, Cheng CS, Hong C, Doty KR, et al. A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* (2009) **138**:114–28. doi:10.1016/j.cell.2009.04.020
- Singh H. Teeing up transcription on CpG islands. *Cell* (2009) **138**:14–6. doi:10.1016/j.cell.2009.06.028
- Ramirez-Carrozzi VR, Nazarian AA, Li CC, Gore SL, Sridharan R, Imbalzano AN, et al. Selective and antagonistic functions of SWI/SNF and Mi-2beta nucleosome remodeling complexes during an inflammatory response. *Genes Dev* (2006) **20**:282–96. doi:10.1101/gad.1383206
- Saunders A, Core LJ, Lis JT. Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* (2006) **7**:557–67. doi:10.1038/nrm1981
- Knoll B, Nordheim A. Functional versatility of transcription factors in the nervous system: the SRF paradigm. *Trends Neurosci* (2009) **32**:432–42. doi:10.1016/j.tins.2009.05.004
- Bhatt DM, Pandya-Jones A, Tong A-J, Barozzi I, Lissner M, Natoli G, et al. Transcript dynamics of pro-inflammatory genes uncovered by RNA-Seq analysis of subcellular RNA fractions. *Cell* (2012) **150**:279–90. doi:10.1016/j.cell.2012.05.043

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 October 2013; accepted: 15 January 2014; published online: 03 February 2014.

Citation: Rudolf R, Busch R, Patra AK, Muhammad K, Avots A, Andrau J-C, Klein-Hessling S and Serfling E (2014) Architecture and expression of the Nfatc1 gene in lymphocytes. *Front. Immunol.* 5:21. doi: 10.3389/fimmu.2014.00021

This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Rudolf, Busch, Patra, Muhammad, Avots, Andrau, Klein-Hessling and Serfling. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Genomic architecture may influence recurrent chromosomal translocation frequency in the *Igh* locus

Amy L. Kenter\*, Robert Wuerffel, Satyendra Kumar and Fernando Grigera

Department of Microbiology and Immunology, University of Illinois College of Medicine, Chicago, IL, USA

\*Correspondence: star1@uic.edu

Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

**Keywords: B cells, *Igh* locus, chromosomal translocations, AID, genomic structure**

## INTRODUCTION

B cell lymphomas represent 95% of all lymphomas diagnosed in the Western world and the majority of these arise from germinal center (GC) B cells (1). Recurrent chromosomal translocations involving Ig loci and proto-oncogenes are a hallmark of many types of B cell lymphoma (2). Three types of breakpoints can be identified in Ig loci. Translocation breakpoints adjacent to the  $D_H$  or  $J_H$  gene segments form secondary to V(D)J recombination, a process that occurs in early B cell development. Other translocations are located in rearranged V(D)J exons that have acquired mutations indicating that translocation is a byproduct of somatic hypermutation (SHM) which occurs in GC B cells. A third type of translocation is characterized by breakpoints in the *Igh* switch regions, a target for double strand DNA breaks (DSBs) during class switch recombination (CSR) that occurs in mature B cells, both inside and outside the GC. Thus, in B lymphocytes, V(D)J joining, CSR, and SHM create obligate single- or double-strand DNA breaks as intermediates for chromosomal translocations (3, 4).

Activation-induced deaminase (AID) is the enzyme that initiates CSR and SHM (5) by inducing the formation of DSBs in switch (S) regions and mutations in V gene exons (6–10). Studies indicate that non-Ig genes are mistargeted by AID (11, 12) and thereby acquire single and double strand DNA breaks at sites coincident with translocation breakpoints (1, 2). Mature B cells are particularly prone to chromosomal translocations that juxtapose Ig genes and proto-oncogenes, including c-myc [Burkitt's lymphoma (BL)], Bcl-2 (follicular lymphoma), Bcl-6 (diffuse large cell lymphoma), and FGFR (multiple myeloma) and which are characteristic of human B cell malignancies (2). The mouse

plasmacytoma (PCT) T(12;15)(*Igh-myc*) translocation, a direct counterpart of the human BL t(8;14)(q24;q32) translocation, occurs as a dynamic process in mature B cells undergoing CSR and is dependent on the expression of AID (13, 14). Hence, a direct mechanistic link between AID and chromosomal translocations focused to Ig genes has been established.

One of the most puzzling aspects of recurrent chromosomal translocations is that DSBs on two different chromosomes must come into close proximity frequently enough to facilitate the crossover. How do the broken ends located at distal sites in cis or on trans chromosomes come together? Consideration of oncogenic selection, sources of translocation prone DSBs associated with antigen receptor rearrangements in B and T lymphocytes, and the role of DSB persistence in translocations have been recently reviewed [(15, 16) and references therein]. Here we consider the proposition that the spatial organization of mammalian genomes is intrinsically linked to genome stability and modulates the frequency of chromosomal translocations.

## A MODEL FOR RECURRENT CHROMOSOMAL TRANSLOCATIONS

Two general models have been proposed to explain the non-random nature of higher order spatial genome organization and the correlation with chromosomal translocations (17). The “contact-first” model posits that translocations require pre-existing physical proximity, whereas, the “breakage-first” model postulates that distant DSBs can be juxtaposed, perhaps through DNA repair machinery. These two theories, the dynamic “breakage-first” and the static “contact-first,” differ fundamentally in their requirement for the presence of DSBs and the mobility of the broken ends.

In the contact-first model only limited local positional motion of DSBs is expected. In the breakage-first model, single DSBs are formed and must undergo large scale movement within nuclei to search for appropriate interaction partners. Although evidence for mobility has been found in yeast systems (18–20), the situation in mammalian cells appears different. In mammalian cells, damaged DNA is largely stationary over time (21–23). However, deprotected telomeres as well as joining of broken DNA ends during V(D)J recombination experience higher mobility (24, 25). Accordingly, the  $V_H$  subdomain of the *Igh* locus has been described as spatially unstructured (26) although additional studies are required to confirm this conclusion. Nevertheless, the weight of evidence in mammalian systems favors the “contact-first” model in light of the limited spatial mobility of DSBs (27). Comparison of a genomic organization map with sites of chromosomal translocation revealed that the spatial proximity of two DSBs is a dominant factor in determining the translocation landscape genome-wide (28). Therefore, it is useful to examine the disposition of loci within chromatin architecture and how this influences the probability of two DSBs finding each other in nuclear space.

## THREE DIMENSIONAL ORGANIZATION OF THE MAMMALIAN GENOME

Emerging evidence indicates that a fundamental property of the mammalian nucleus is the non-random organization of the genome in nuclear space (29). Cytogenetic studies reveal that the mammalian nucleus is occupied by non-randomly positioned genes and chromosomes (30). Together these studies have shown that gene activation or silencing is often associated with repositioning of that locus relative to nuclear compartments and



other genomic loci. In this regard, it is relevant that in normal B cells, the breakage sites of several common translocations are more frequently found in close spatial proximity in the nucleus than would be expected based on random positioning (31). A similar relationship between translocation frequency and spatial proximity is observed in BL where the *myc* locus is on average closest to its most frequent translocation partner, *Igh* (32). The non-random aspect of genome spatial organization in a sub-compartmentalized nuclear space has emerged as a potential contributor to the genesis of chromosomal translocations (23).

The combination of new imaging tools and the comprehensive mapping of long range chromosomal interaction has revealed structural features and biological properties of the three dimensional (3D) genomic organization (33–38). Four features contributing to an ordered 3D organization of eukaryotic genomes have become evident. (1) Individual chromosomes occupy distinct chromosomal territories (CT) with only a limited degree of intermingling (39). (2) The eukaryotic genome is partitioned into functionally distinct euchromatin and heterochromatin (40). (3) Individual genomic loci and elements display preferences for nuclear positioning which correlates well with genomic functions including transcriptional activity and replication timing (39, 41). (4) Distant chromosomal elements associate to form chromatin loops thereby providing a mechanism for long range enhancer function (36, 38, 42). These variables predict that unique and unanticipated spatial genomic relationships may determine unique combinations of chromosomal translocations that may differ in specific tissues and during differentiation.

### CHROMOSOMAL LOOPING INTERACTIONS FACILITATE CSR

The best studied property of chromatin looping is the spatial proximity of genes and their regulatory elements to establish functional states. Of relevance here is the recognition that chromatin looping influences partner selection during V(D)J recombination (43–45), CSR (46, 47), and may drive specific chromosomal translocation events (28, 48, 49). It is of importance to understand the spatial relationships

within the *Igh* locus and how they relate to the preferential expression of Ig gene expression and protect against genome instability. We focus here on CSR because the most prevalent B cell lymphomas arise from GC B cells and are dependent on the expression of AID (1, 13, 14).

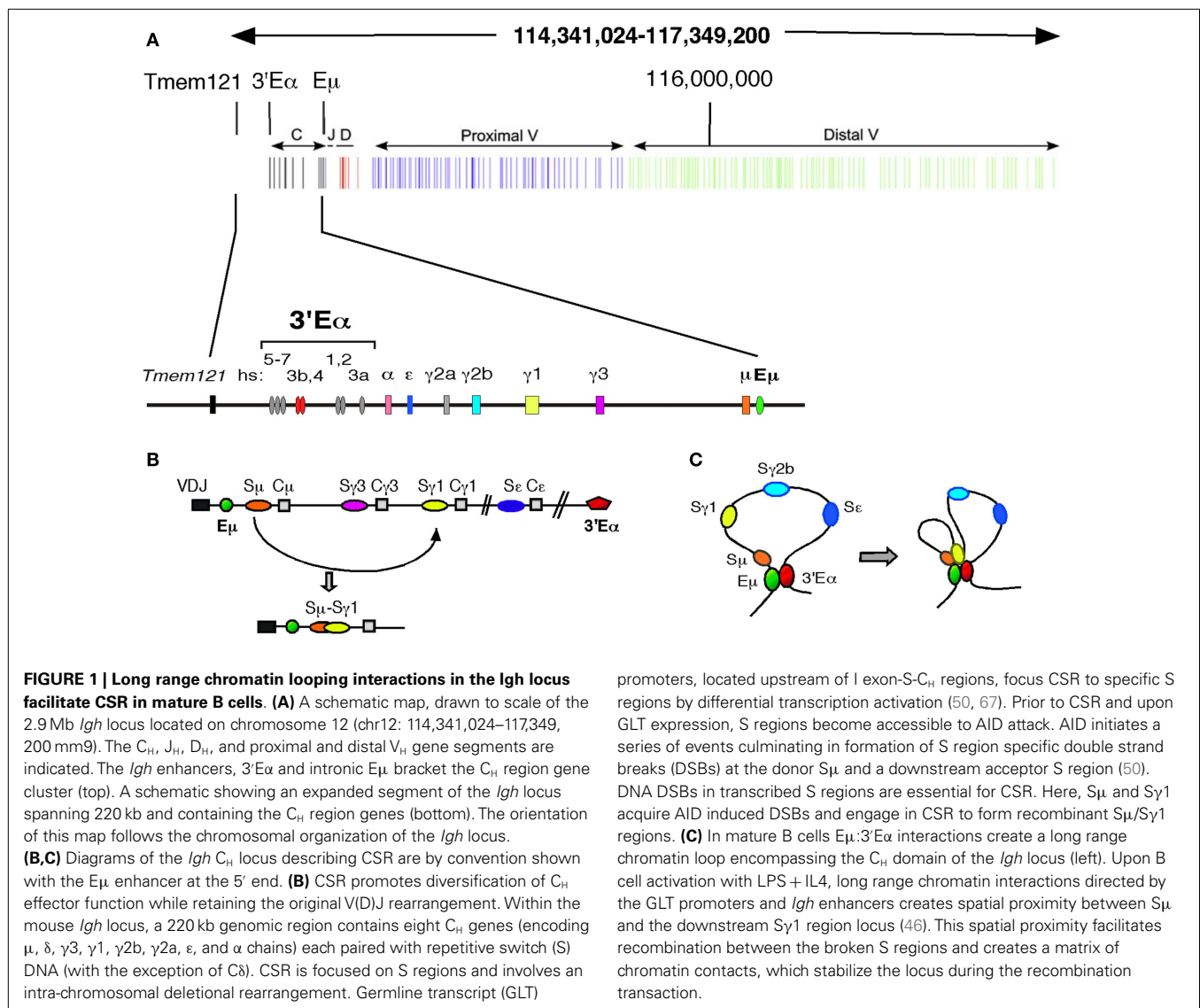
Class switch recombination promotes diversification of  $C_H$  effector function while retaining the original rearranged V(D)J exons. The mouse *Igh* locus spans 2.9 Mb within which a centromeric 220 kb genomic region contains eight  $C_H$  genes (encoding  $\mu$ ,  $\delta$ ,  $\gamma 3$ ,  $\gamma 1$ ,  $\gamma 2b$ ,  $\gamma 2a$ ,  $\epsilon$ , and  $\alpha$  chains) each paired with repetitive S DNA (with the exception of  $C\delta$ ) (Figure 1A). CSR is focused on S regions and involves an intra-chromosomal deletional rearrangement (Figure 1B). Germline transcript (GLT) promoters, located upstream of I exon-S- $C_H$  regions, focus CSR to specific S regions by differential transcription activation (9, 50). The I-S- $C_H$  region genes are embedded between the  $E\mu$  intronic and 3'E $\alpha$  enhancers (51). Chromosome conformation capture (3C) studies reveal that in mature resting B cells the transcriptional enhancer elements,  $E\mu$  and 3'E $\alpha$ , engage in long range chromatin looping interactions (46, 47) (Figure 1C). B cell activation leads to induced recruitment of the GLT promoters to the  $E\mu$ :3'E $\alpha$  complex that in turn facilitates GLT expression and supports S/S synapsis (46).

The 3'E $\alpha$  regulatory region plays a significant role in mediating the spatial structure of the *Igh* locus during CSR as well as promoting genome stability (52). Targeted deletion of *hs3b*, 4 within 3'E $\alpha$  abolishes GLT expression and GLT promoter:3'E $\alpha$  and  $E\mu$ :3'E $\alpha$  looping interactions (46, 53, 54). AID initiates a series of events ending in creation of S region specific DNA DSBs at the donor  $S\mu$  and a downstream acceptor S region to create S/S junctions and facilitate CSR (7). S regions targeted by AID for DSB formation are transcriptionally active. Chromatin looping across this region ensures proximity between two S regions targeted for DSB creation and recombination (Figure 1C). Thus, CSR is dependent on 3D chromatin architecture mediated by long range intra-chromosomal interactions between distantly located transcriptional elements that serves to tether broken chromosomal DNA together during the CSR reaction.

Chromosome conformation capture (3C, 4C, 5C, and Hi-C) based studies indicate that the most probable chromatin interactions are the most proximal ones and the probability of contact decreases with distance. Correspondingly, alignment of genomic organization maps with sites of chromosomal translocation generated in Hi-C and 4C studies have shown that translocations are enriched in cis along single chromosomes containing the target DSB and in trans in a manner related to pre-existing spatial proximity (28, 55). The positional immobilization of DSBs in the *Igh* locus, for example, should render the probability of successful translocation as the product of the frequency of each DSB at the sites of crossover and the frequency with which these sites are synapsed in physical space (28). In B lymphocytes *c-myc/Igh* translocations occur in trans and may represent a failure of stringent spatial sequestration of AID induced DSBs to within the *Igh* locus (56, 57).

### DYNAMIC CHROMATIN INTERACTIONS AND THE GENESIS OF CHROMOSOMAL TRANSLOCATIONS

Chromosomal translocation frequency as reported by genome-wide translocation sequencing is determined by the frequency of AID induced DSB at translocation targets, factors that contribute to synapsis of broken loci, and circumvention of DNA repair functions that facilitate intra-chromosomal DSB joining (55–58). Are recurrent chromosomal translocations simply the result of a stochastic process related to the probability of contact between AID induced DSBs? Tagging single loci with Lac operon (LacO) arrays, as well as photobleaching and photoactivation experiments, have shown that interphase chromatin is locally mobile but rarely moves over long distances (59–61). However, lamina associated domains are large genomic regions that are in intermittent molecular contact with the nuclear lamina indicating a dynamic spatial architecture of chromosomes (62). Chromatin looping, clustering, and compartmentalization are dynamic and responsive to developmental and environmental cues. Functionally dynamic chromatin responses include formation of transcription and replication factories, and nuclear relocation of loci during development (63–66). The looping



interactions spanning the *IgH* locus during CSR and in the presence of DSBs may also be dynamic and to some degree transient. In a dynamic chromosomal setting, DSBs present in an *IgH* locus that lacks *Eμ*:3'Eα tethering, for example, would be at high risk of re-joining to sites outside the *IgH* locus along chromosome 12 and at lower frequency to sites on other chromosomes. The dynamism of chromosomal transactions are not yet fully described and represent the next forefront for investigation to appreciate constraints and variables of genome stability and instability.

## AUTHOR CONTRIBUTIONS

Drs. Robert Wuerffel, Satyendra Kumar, Fernando Grigera, and Amy L. Kenter were

all involved in developing the ideas regarding long range chromatin interactions and dynamics that are the subject here and all have critiqued and agree to the contents of this piece. Amy L. Kenter wrote the article.

## ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health RO1AI052400 and R21AI106328 to Amy L. Kenter.

## REFERENCES

1. Kuppers R. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer* (2005) 5:251–62. doi:10.1038/nrc1589
2. Kuppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* (2001) 20:5580–94. doi:10.1038/sj.onc.1204640

3. Nussenzweig A, Nussenzweig MC. Origin of chromosomal translocations in lymphoid cancer. *Cell* (2010) 141:27–38. doi:10.1016/j.cell.2010.03.016
4. Tsai AG, Lu H, Raghavan SC, Muschen M, Hsieh CL, Lieber MR. Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. *Cell* (2008) 135:1130–42. doi:10.1016/j.cell.2008.10.035
5. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* (2000) 102:553–63. doi:10.1016/S0092-8674(00)00078-7
6. Chahwan R, Edelman W, Scharff MD, Roa S. AID-ing antibody diversity by error-prone mismatch repair. *Semin Immunol* (2012) 24:293–300. doi:10.1016/j.smim.2012.05.005
7. Kenter AL. AID targeting is dependent on RNA polymerase II pausing. *Semin Immunol* (2012) 24:281–6. doi:10.1016/j.smim.2012.06.001

8. Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, et al. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* (2010) **143**:122–33. doi:10.1016/j.cell.2010.09.017
9. Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. *Annu Rev Immunol* (2008) **26**:261–92. doi:10.1146/annurev.immunol.26.021607.090248
10. Peled JU, Kuang FL, Iglesias-Ussel MD, Roa S, Kalis SL, Goodman MF, et al. The biochemistry of somatic hypermutation. *Annu Rev Immunol* (2008) **26**:481–511. doi:10.1146/annurev.immunol.26.021607.090236
11. Shen HM, Peters A, Baron B, Zhu X, Storb U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science* (1998) **280**:1750–2. doi:10.1126/science.280.5370.1750
12. Kotani AI, Okazaki M, Muramatsu M, Kinoshita K, Begum NA, Nakajima T, et al. A target selection of somatic hypermutations is regulated similarly between T and B cells upon activation-induced cytidine deaminase expression. *Proc Natl Acad Sci U S A* (2005) **102**:4506–11. doi:10.1073/pnas.0500830102
13. Ramiro AR, Jankovic M, Callen E, Difilippantonio S, Chen HT, McBride KM, et al. Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. *Nature* (2006) **440**:105–9. doi:10.1038/nature04495
14. Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, et al. AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* (2004) **118**:431–8. doi:10.1016/j.cell.2004.08.006
15. Alt FW, Zhang Y, Meng FL, Guo C, Schwer B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* (2013) **152**:417–29. doi:10.1016/j.cell.2013.01.007
16. Gostissa M, Alt FW, Chiarle R. Mechanisms that promote and suppress chromosomal translocations in lymphocytes. *Annu Rev Immunol* (2011) **29**:319–50. doi:10.1146/annurev-immunol-031210-101329
17. Misteli T, Soutoglou E. The emerging role of nuclear architecture in DNA repair and genome maintenance. *Nat Rev Mol Cell Biol* (2009) **10**:243–54. doi:10.1038/nrm2651
18. Lisby M, Antunez de Mayolo A, Mortensen UH, Rothstein R. Cell cycle-regulated centers of DNA double-strand break repair. *Cell Cycle* (2003) **2**:479–83. doi:10.4161/cc.2.5.483
19. Lisby M, Mortensen UH, Rothstein R. Colocalization of multiple DNA double-strand breaks at a single Rad52 repair centre. *Nat Cell Biol* (2003) **5**:572–7. doi:10.1038/ncb997
20. Lisby M, Rothstein R. DNA damage checkpoint and repair centers. *Curr Opin Cell Biol* (2004) **16**:328–34. doi:10.1016/j.ccb.2004.03.011
21. Nelms BE, Maser RS, MacKay JF, Lagally MG, Petrini JH. In situ visualization of DNA double-strand break repair in human fibroblasts. *Science* (1998) **280**:590–2. doi:10.1126/science.280.5363.590
22. Kruhlak MJ, Celeste A, Nussenzweig A. Spatio-temporal dynamics of chromatin containing DNA breaks. *Cell Cycle* (2006) **5**:1910–2. doi:10.4161/cc.5.17.3169
23. Meaburn KJ, Misteli T, Soutoglou E. Spatial genome organization in the formation of chromosomal translocations. *Semin Cancer Biol* (2007) **17**:80–90. doi:10.1016/j.semcancer.2006.10.008
24. Dimitrova N, Chen YC, Spector DL, de Lange T. 53BP1 promotes non-homologous end joining of telomeres by increasing chromatin mobility. *Nature* (2008) **456**:524–8. doi:10.1038/nature07433
25. Difilippantonio S, Gapud E, Wong N, Huang CY, Mahowald G, Chen HT, et al. 53BP1 facilitates long-range DNA end-joining during V(D)J recombination. *Nature* (2008) **456**:529–33. doi:10.1038/nature07476
26. Medvedovic J, Ebert A, Tagoh H, Tamir IM, Schwickert TA, Novatchkova M, et al. Flexible long-range loops in the VH gene region of the Igh locus facilitate the generation of a diverse antibody repertoire. *Immunity* (2013) **39**:229–44. doi:10.1016/j.immuni.2013.08.011
27. Soutoglou E, Dorn JF, Sengupta K, Jasim M, Nussenzweig A, Ried T, et al. Positional stability of single double-strand breaks in mammalian cells. *Nat Cell Biol* (2007) **9**:675–82. doi:10.1038/ncb1591
28. Zhang Y, McCord RP, Ho YJ, Lajoie BR, Hildebrand DG, Simon AC, et al. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* (2012) **148**:908–21. doi:10.1016/j.cell.2012.02.002
29. Lancot C, Cheutin T, Cremer M, Cavalli G, Cremer T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* (2007) **8**:104–15. doi:10.1038/nrg2041
30. Meaburn KJ, Misteli T. Cell biology: chromosome territories. *Nature* (2007) **445**:379–781. doi:10.1038/445379a
31. Neves H, Ramos C, da Silva MG, Parreira A, Parreira L. The nuclear topography of ABL, BCR, PML, and RARalpha genes: evidence for gene proximity in specific phases of the cell cycle and stages of hematopoietic differentiation. *Blood* (1999) **93**:1197–207.
32. Roix JJ, McQueen PG, Munson PJ, Parada LA, Misteli T. Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat Genet* (2003) **34**:287–91. doi:10.1038/ng1177
33. Baker M. Genomics: genomes in three dimensions. *Nature* (2011) **470**:289–94. doi:10.1038/470289a
34. Osborne CS, Ewels PA, Young AN. Meet the neighbours: tools to dissect nuclear structure and function. *Brief Funct Genomics* (2011) **10**:11–7. doi:10.1093/bfpg/elq034
35. van Steensel B, Dekker J. Genomics tools for unraveling chromosome architecture. *Nat Biotechnol* (2010) **28**:1089–95. doi:10.1038/nbt.1680
36. Naumova N, Dekker J. Integrating one-dimensional and three-dimensional maps of genomes. *J Cell Sci* (2010) **123**:1979–88. doi:10.1242/jcs.051631
37. Cavalli G, Misteli T. Functional implications of genome topology. *Nat Struct Mol Biol* (2013) **20**:290–9. doi:10.1038/nsmb.2474
38. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* (2012) **485**:376–80. doi:10.1038/nature11082
39. Cremer T, Cremer M. Chromosome territories. *Cold Spring Harb Perspect Biol* (2010) **2**:a003889. doi:10.1101/cshperspect.a003889
40. Felsenfeld G, Groudine M. Controlling the double helix. *Nature* (2003) **421**:448–53. doi:10.1038/nature01411
41. Takizawa T, Meaburn KJ, Misteli T. The meaning of gene positioning. *Cell* (2008) **135**:9–13. doi:10.1016/j.cell.2008.09.026
42. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragozy T, Telling A, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* (2009) **326**:289–93. doi:10.1126/science.1181369
43. Guo C, Gerasimova T, Hao H, Ivanova I, Chakraborty T, Selimyan R, et al. Two forms of loops generate the chromatin conformation of the immunoglobulin heavy-chain gene locus. *Cell* (2011) **147**:332–43. doi:10.1016/j.cell.2011.08.049
44. Guo C, Yoon HS, Franklin A, Jain S, Ebert A, Cheng HL, et al. CTCF-binding elements mediate control of V(D)J recombination. *Nature* (2011) **477**:424–30. doi:10.1038/nature10495
45. Jhunjhunwala S, van Zelm MC, Peak MM, Murre C. Chromatin architecture and the generation of antigen receptor diversity. *Cell* (2009) **138**:435–48. doi:10.1016/j.cell.2009.07.016
46. Wuerffel R, Wang L, Grigera F, Manis J, Selsing E, Perlot T, et al. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity* (2007) **27**:711–22. doi:10.1016/j.immuni.2007.09.007
47. Sellars M, Reina-San-Martin B, Kastner P, Chan S. Ikaros controls isotype selection during immunoglobulin class switch recombination. *J Exp Med* (2009) **206**:1073–87. doi:10.1084/jem.20082311
48. Lin C, Yang L, Tanasa B, Hutt K, Ju BG, Ohgi K, et al. Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* (2009) **139**:1069–83. doi:10.1016/j.cell.2009.11.030
49. Mani RS, Tomlins SA, Callahan K, Ghosh A, Nyati MK, Varambally S, et al. Induced chromosomal proximity and gene fusions in prostate cancer. *Science* (2009) **326**:1230. doi:10.1126/science.1178124
50. Chaudhuri J, Basu U, Zarrin A, Yan C, Franco S, Perlot T, et al. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv Immunol* (2007) **94**:157–214. doi:10.1016/S0065-2776(06)94006-1
51. Perlot T, Alt FW. Cis-regulatory elements and epigenetic changes control genomic rearrangements of the IgH locus. *Adv Immunol* (2008) **99**:1–32. doi:10.1016/S0065-2776(08)00601-9
52. Gostissa M, Yan CT, Bianco JM, Cogne M, Pinaud E, Alt FW. Long-range oncogenic activation of Igh-c-myc translocations by the Igh 3' regulatory region. *Nature* (2009) **462**:803–7. doi:10.1038/nature08633
53. 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**:187–99. doi:10.1016/S1074-7613(01)00181-9
54. Pinaud E, Marquet M, Fiancette R, Peron 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
  55. Rocha PP, Micsinai M, Kim JR, Hewitt SL, Souza PP, Trimarchi T, et al. Close proximity to Igh is a contributing factor to AID-mediated translocations. *Mol Cell* (2012) **47**:873–85. doi:10.1016/j.molcel.2012.06.036
  56. Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* (2011) **147**:107–19. doi:10.1016/j.cell.2011.07.049
  57. Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, et al. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell* (2011) **147**:95–106. doi:10.1016/j.cell.2011.07.048
  58. Hakim O, Resch W, Yamane A, Klein I, Kieffer-Kwon KR, Jankovic M, et al. DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. *Nature* (2012) **484**:69–74. doi:10.1038/nature10909
  59. Walter J, Schermelleh L, Cremer M, Tashiro S, Cremer T. Chromosome order in HeLa cells changes during mitosis and early G1, but is stably maintained during subsequent interphase stages. *J Cell Biol* (2003) **160**:685–97. doi:10.1083/jcb.200211103
  60. Strickfaden H, Zunhammer A, van Koningsbruggen S, Kohler D, Cremer T. 4D chromatin dynamics in cycling cells: Theodor Boveri's hypotheses revisited. *Nucleus* (2010) **1**:284–97. doi:10.4161/nucl.1.3.11969
  61. Thomson I, Gilchrist S, Bickmore WA, Chubb JR. The radial positioning of chromatin is not inherited through mitosis but is established de novo in early G1. *Curr Biol* (2004) **14**:166–72. doi:10.1016/j.cub.2003.12.024
  62. Kind J, Pagie L, Ortaobokoyun H, Boyle S, de Vries SS, Janssen H, et al. Single-cell dynamics of genome-nuclear lamina interactions. *Cell* (2013) **153**:178–92. doi:10.1016/j.cell.2013.02.028
  63. Rajapakse I, Perlman MD, Scalzo D, Kooperberg C, Groudine M, Kosak ST. The emergence of lineage-specific chromosomal topologies from coordinate gene regulation. *Proc Natl Acad Sci U S A* (2009) **106**:6679–84. doi:10.1073/pnas.0900986106
  64. Peric-Hupkes D, Meuleman W, Pagie L, Brugge-man SW, Solovei I, Brugman W, et al. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell* (2010) **38**:603–13. doi:10.1016/j.molcel.2010.03.016
  65. Edelman LB, Fraser P. Transcription factories: genetic programming in three dimensions. *Curr Opin Genet Dev* (2012) **22**:110–4. doi:10.1016/j.gde.2012.01.010
  66. Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* (2002) **296**:158–62. doi:10.1126/science.1068768
  67. Stavnezer J. Molecular processes that regulate class switching. *Curr Top Microbiol Immunol* (2000) **245**:127–68.

Received: 03 December 2013; accepted: 18 December 2013; published online: 30 December 2013.

Citation: Kenter AL, Wuerffel R, Kumar S and Grigera F (2013) Genomic architecture may influence recurrent chromosomal translocation frequency in the IgH locus. *Front. Immunol.* 4:500. doi: 10.3389/fimmu.2013.00500  
This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2013 Kenter, Wuerffel, Kumar and Grigera. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# AIDing chromatin and transcription-coupled orchestration of immunoglobulin class-switch recombination

Bharat Vaidyanathan<sup>1,2†</sup>, Wei-Feng Yen<sup>1,2†</sup>, Joseph N. Pucella<sup>2†</sup> and Jayanta Chaudhuri<sup>1,2\*</sup>

<sup>1</sup> Weill Cornell Graduate School of Medical Sciences, New York, NY, USA

<sup>2</sup> Immunology Program, Memorial Sloan Kettering Cancer Center, Gerstner Sloan Kettering Graduate School, New York, NY, USA

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

Amy L. Kenter, University of Illinois College of Medicine, USA  
Sebastian Dudo Fugmann, Chang Gung University, Taiwan

## \*Correspondence:

Jayanta Chaudhuri, Immunology Program, Memorial Sloan Kettering Cancer Center, Gerstner Sloan Kettering Graduate School, New York, NY 10065, USA  
e-mail: [chaudhuri@mskcc.org](mailto:chaudhuri@mskcc.org)

<sup>†</sup> Bharat Vaidyanathan, Wei-Feng Yen and Joseph N. Pucella have contributed equally to this work.

Secondary diversification of the antibody repertoire upon antigenic challenge, in the form of immunoglobulin heavy chain (IgH) class-switch recombination (CSR) endows mature, naïve B cells in peripheral lymphoid organs with a limitless ability to mount an optimal humoral immune response, thus expediting pathogen elimination. CSR replaces the default constant ( $C_H$ ) region exons ( $C_\mu$ ) of IgH with any of the downstream  $C_H$  exons ( $C_\gamma$ ,  $C_\epsilon$ , or  $C_\alpha$ ), thereby altering effector functions of the antibody molecule. This process depends on, and is orchestrated by, activation-induced deaminase (AID), a DNA cytidine deaminase that acts on single-stranded DNA exposed during transcription of switch (S) region sequences at the IgH locus. DNA lesions thus generated are processed by components of several general DNA repair pathways to drive CSR. Given that AID can instigate DNA lesions and genomic instability, stringent checks are imposed that constrain and restrict its mutagenic potential. In this review, we will discuss how AID expression and substrate specificity and activity is rigorously enforced at the transcriptional, post-transcriptional, post-translational, and epigenetic levels, and how the DNA-damage response is choreographed with precision to permit targeted activity while limiting bystander catastrophe.

**Keywords:** cytidine deamination, DNA recombination, DNA repair, class-switching, R-loops

## INTRODUCTION

B cells are specialized lymphocytes that express Ig receptors (or antibodies) on their cell surface. Antibodies are comprised of immunoglobulin heavy chains (IgH) and light chains (IgL), with the N-termini of IgH and IgL generating the antigen-binding pocket, and the C-terminus of IgH performing effector functions. A salient feature of B-lymphocytes is their ability to recognize an almost infinite array of antigens. This enormous diversity is achieved through V(D)J recombination, a process that assembles the exons encoding the amino-terminal variable regions of IgH and IgL from component variable (V), diversity (D), and joining (J) segments (1). The end product of V(D)J recombination is a mature but naïve  $IgM^+$  B cell that exits the bone marrow. In the context of specialized structures called germinal centers in secondary lymphoid organs such as the spleen and lymph nodes, mature B cells interact with antigens and undergo class-switch recombination (CSR) (2, 3).

The mouse IgH locus is comprised of eight constant region ( $C_H$ ) exons, with  $C_\mu$  most proximal to the variable region segments and  $C_\alpha$  being the most distal (Figure 1). CSR exchanges the default  $C_\mu$  for an alternative set of downstream  $C_H$  exons, for example,  $C_\gamma$ ,  $C_\epsilon$ , or  $C_\alpha$ , so that the B cell switches from expressing IgM to one producing a secondary antibody isotype such as IgG, IgE, or IgA, respectively. CSR occurs between repetitive “switch” (S) DNA elements that precede each set of  $C_H$  exons. According to the conventional model for CSR, transcription through S regions promotes formation of DNA:RNA hybrid structures, such as R-loops that reveal single-stranded DNA (ssDNA) substrates for activation-induced deaminase (AID)-mediated cytidine

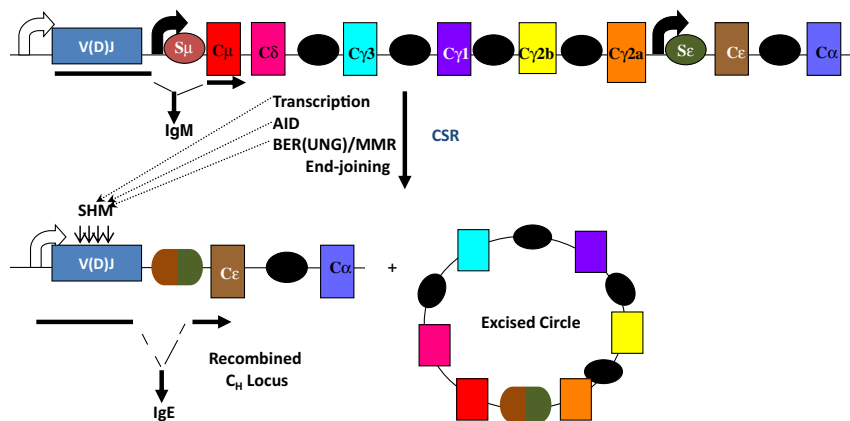
deamination (Figure 2). The deaminated residues are processed into DNA double-strand breaks (DSBs) by components of the base-excision repair (BER) and mismatch repair (MMR) pathways (4–6). End-joining of DSBs between two S regions results in the excision of the intervening sequence and juxtaposition of a new set of constant region exons directly downstream of the rearranged V(D)J segment, thereby generating Ig molecules with the same antigen specificity but with new effector functions (2, 3) (Figure 1). Here, we will discuss the intrinsic properties of AID, the factors commissioning its function to produce obligatory DNA DSBs intermediates, and the DNA repair/end-joining pathways that ensure productive recombination.

Germinal center B cells also undergo another AID-dependent secondary diversification reaction termed somatic hypermutation (SHM) wherein point mutations, and sometimes insertions and deletions, are introduced at a very high rate ( $10^{-2}$ – $10^{-3}$ /bp/generation) into the recombined variable region exons encoding IgH and IgL, so as to select B cells with increased antigen affinity (Figure 1). SHM requires transcription through the variable region exons and occurs primarily, but not exclusively, at RGYW “hot-spot” motifs where R = purine base, Y = pyrimidine base, and W = A or T nucleotide. Details of SHM are outside the scope of this review and have been discussed in multiple excellent reviews [for example, Ref. (7)].

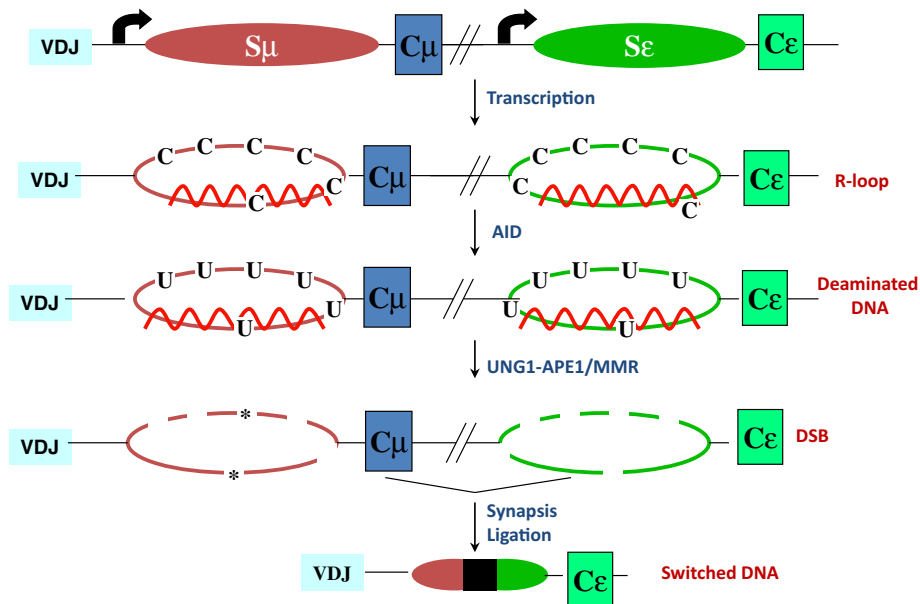
## INTRINSIC PROPERTIES OF AID

Following the discovery of AID by subtractive cDNA hybridization from the mouse CH12F3 B lymphoma cells and its proven essentiality for SHM and CSR (8, 9), a huge amount of effort has





**FIGURE 1 | Overview of CSR.** CSR is a deletional-recombination reaction between repetitive switch (S) regions (ovals) that precede each set of constant region ( $C_H$ ) exons. Cytokine-induced transcription through S regions drives CSR to specific  $C_H$  exons. CSR requires AID and components of the base-excision repair (BER) and mismatch repair (MMR) pathways. Somatic hypermutation (SHM) introduces point mutations at a very high rate into variable region exons, V(D)J.



**FIGURE 2 | R-loop-based model for CSR.** Transcription through S regions generates R-loop structures in which the RNA stably hybridizes to the template strand, displacing the non-template strand as ssDNA. AID activity at R-loops generate uridines in the DNA, which are subsequently processed by BER and MMR proteins to ultimately generate DSBs. End-joining of DSBs completes CSR. Asterisks represent nicks close to each other on opposite strands.

gone into the characterization of its enzymatic properties. Elegant biochemical work using purified AID from activated splenic B cells, recombinant GST-AID from Sf9 cells, and other epitope-tagged forms, has shed light into the DNA deamination ability of AID *in vitro* [reviewed in Ref. (2)]. These studies demonstrated unequivocally that AID deaminates deoxycytidines (dCs) in ssDNA, and fails to act on dsDNA, RNA, and DNA:RNA hybrids. Additionally, it was shown that AID could deaminate dCs in the context of transcribed dsDNA (10, 11), suggesting that access to and activity on *in vivo* substrates might require transcription of the

locus. Since the crystal structure of AID has not been determined, the field has faced a bottleneck in explicit elucidation of enzyme biochemistry. Still, based on structural and biochemical insights from bacterial cytidine deaminases and related DNA/RNA deaminases such as APOBECs, the mechanism of  $Zn^{2+}$ -dependent catalysis by the active site (H56, E58, C87, C90) residues and preference for RGYW motif (residues 113–123) was cogently demonstrated (12, 13).

*In vitro* deamination assays using recombinant GST-AID purified from insect cells suggest that AID performs processive catalysis

(14), which leads to accumulation of multiple mutations on a single DNA fragment, disfavoring “jumping” onto a second fragment. This finding is in contrast to proposed distributive mode of action based on the high net positive charge (+11 at pH 7.0) of AID that promotes strong binding to nucleic acids (15). Nonetheless, *in vitro* deamination assays performed on ssDNA substrates revealed that AID-mediated deamination is intrinsically inefficient, haphazard, and a “random bidirectional walk” along DNA, yielding ~3% deamination upon hot-spot encounter (16). Such a mechanism has likely evolved to generate a diverse array of mutations, especially to favor the selection of high affinity antibodies *in vivo* (16). A note of caution to be borne in mind while interpreting these biochemical analyses is that the bulky tag might affect inherent properties of AID, and GST-AID does not reconstitute CSR *in vivo* (17), suggesting that *in vitro* results may not accurately reflect the *in vivo* scenario. Besides, this form of AID requires the action of RNase A to be active *in vitro*, which contradicts reports of AID purified from B cells (10), advocating for adventitious properties unique to GST-AID. It is to be noted that AID, based on its homology to the RNA editing enzyme APOBEC1, has been proposed to edit mRNAs and/or micro-RNAs (miRs) required for CSR and SHM, however, there is no experimental evidence yet to support this notion (18). Thus, despite the limitations of AID enzyme biology, strong genetic evidence has propelled DNA deamination as being the generally accepted model, and this review is based on this premise.

### REGULATION OF AID EXPRESSION, LOCALIZATION, AND STABILITY

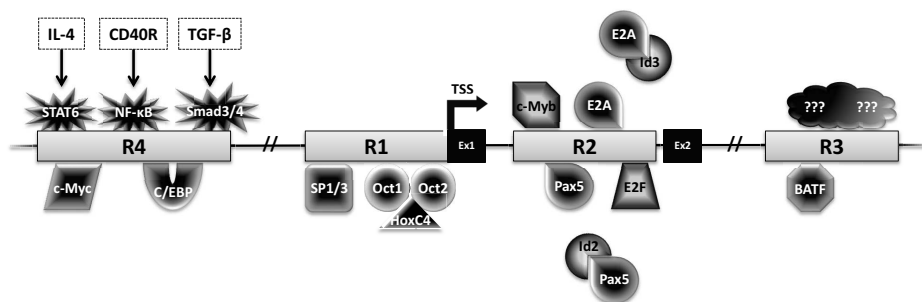
While the primary and physiological role of AID is to introduce DNA lesions at the Ig loci to drive antibody diversification, AID also poses a threat to genomic integrity. Ectopic expression of AID in non-B cells converts it into a general mutator (19, 20). Even in B cells, mistargeted AID activity is the major underlying cause behind oncogenic translocations that are hallmarks of a large number of B cell malignancies (1, 21). Therefore, regulation of AID expression is fundamental not only for the development of an efficient immune system, but also for the maintenance of genomic integrity inherent to cells expressing a mutator. Thus, it is not surprising that AID comes outfitted with multifaceted transcriptional and post-transcriptional regulatory mechanisms.

### TRANSCRIPTIONAL REGULATION OF AID EXPRESSION

Activation-induced deaminase is encoded by the *Aicda* gene, located on chromosome 6 and 12 in mice and humans, respectively. Four highly conserved regulatory regions activate *Aicda* transcription primarily in activated B cells, and restrict its expression in other cell types (Figure 3). Region 1 is comprised of a TATA-less promoter and enhancer elements that bind HoxC4–Oct1/2 and Sp1/3 [reviewed in Ref. (22)]. This region also contains elements that respond to estrogen and progesterone, hormones that, respectively, activate or repress AID expression (23–25). Region 2 lies within the first *Aicda* intron and includes binding sites for B-cell-specific Pax5 and E2A proteins (22). This region also harbors silencer elements that could bind repressors E2F and c-Myb in a fashion unrestricted to B cells (22). Deletion of the silencer elements drastically increases AID expression, without inducing transcription in non-B cells, bolstering the notion of extensive checks to AID expression (26). Region 3, approximately 25 kb downstream of *Aicda*, is necessary to sustain physiological levels of AID expression, likely through a BATF-binding site (22, 27, 28). Region 4 is approximately 8 kb upstream of the *Aicda* transcriptional initiation site and contains enhancers that bind NF-κB, STAT6, and SMAD3/4, factors that are stimulated by B cell activation (22). Recently, c-Myc was implicated in binding Region 4 to promote robust AID expression (29, 30).

Although physiological AID expression is largely restricted to mature B cells, its expression has also been reported in other settings. AID is expressed in developing B cells in the bone marrow, inducing robust CSR to a subset of isotypes (31, 32). The physiological relevance of CSR in the bone marrow is not clear at present. AID expression has also been observed in intestinal epithelial cells during *Helicobacter pylori* infection; whether this represents aberrant expression or some uncharacterized response to infection is not known (33, 34). Additionally, AID expression has been observed in prostate cancer cells (35); such aberrant AID expression might be correlative or causal to pathological outcomes.

It is not clear at present whether the non-B-cell-specific expression of AID has any physiological relevance, and the AID fate-mapping mouse does not reveal robust expression pattern in non-lymphoid cells (36). But an intriguing finding is that AID is expressed in primordial germ cells, in embryonic stem (ES) cells, and also in mouse embryonic fibroblasts induced to



**FIGURE 3 | Transcriptional regulation of AID.** The *Aicda* locus contains four conserved regulatory regions (R1–4). The first two exons (Ex1–2) and transcription factors with the potential to bind these regions are shown. Factors with black center and white text are activators, while factors with white center and black text are repressors.

undergo transcription-factor-mediated reprogramming (37, 38). In this regard, AID has been posited to deaminate methylated cytidine, and in concert with DNA BER, promote demethylation of genes required for the maintenance of a pluripotent stem-cell state (37–40). However, AID has extremely weak intrinsic activity on 5mC, and AID-deficient mice do not exhibit any overt phenotype or methylome changes that could be attributed to a failure in active demethylation (41, 42). Thus, *in vivo* demethylation by AID and the factors regulating AID expression in these settings remain provocative and warrant further research.

#### MICRO-RNA-MEDIATED REGULATION OF AID EXPRESSION

Another level of regulation exists at the level of stability of *Aicda* mRNA, enforced by miRs such as miR-155, miR-181b, and miR-361, with miR-155 being the best characterized. miR-155 expression is upregulated upon activation for CSR. The 3'-untranslated region (3'-UTR) of *Aicda* mRNA has a binding site for miR-155, mutation of which, increased AID expression and doubled the frequency of CSR (43, 44). Surprisingly, although miR-155-deficient B cells upregulate AID expression, they do not undergo increased CSR, perhaps due to dysregulation of other miR-155 targets relevant to CSR (45). The 3'-UTR of *Aicda* mRNA also contains a binding site for miR-361 (46). Significantly, the transcription factor Bcl6, required for formation of germinal centers, binds and transcriptionally represses both miR-155 and miR-361, in turn relieving repression of AID (46). The role, if any, of miR-361 in the regulation of AID mRNA stability remains to be determined. Finally, ectopic expression of miR-181b in activated murine B cells impaired CSR, likely due to reduced AID mRNA and protein levels (47). Given the emerging significance of canonical and non-canonical miR targeting, it can be conceived that many more miRs affecting AID and CSR are awaiting discovery (48).

#### SUBCELLULAR LOCALIZATION AND STABILITY

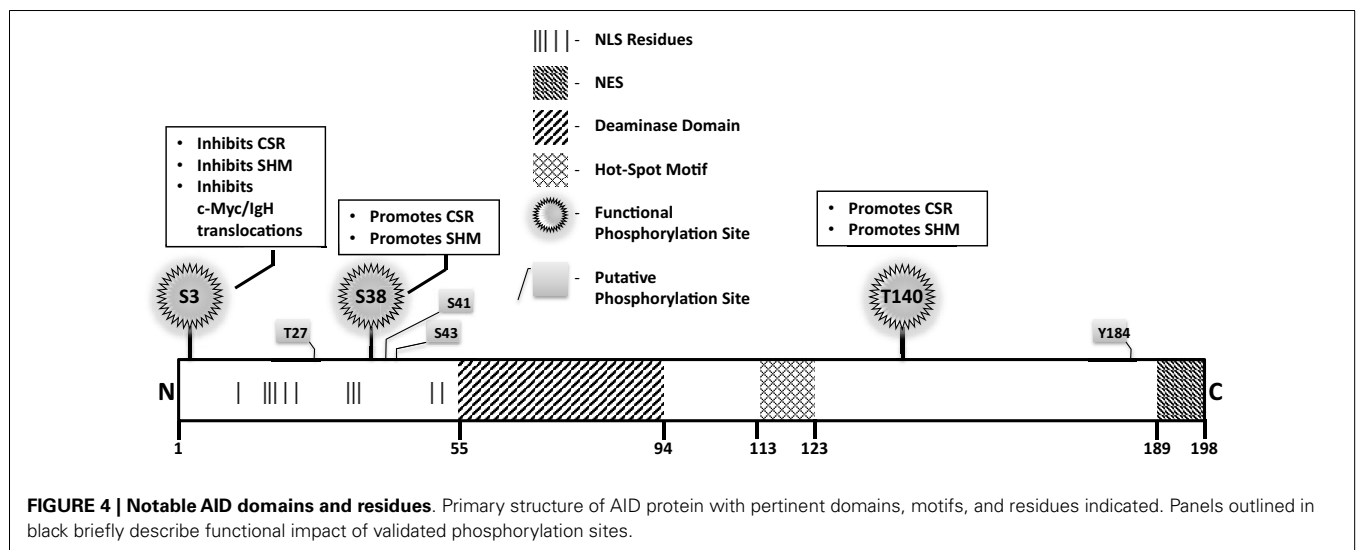
A rational way to constrain AID activity on DNA is by regulating its subcellular localization. AID localization is governed by active nuclear import, cytoplasmic retention, and efficient nuclear export. The majority (greater than 90%) of AID is sequestered in

the cytoplasm, possibly through interactions of the C-terminus of AID with eEF1A, chaperone Hsp90, and co-chaperone Hsp40 DnaJ1 (49). Nuclear entry of AID is dependent on importin-3 and a conformational nuclear localization signal (NLS) generated upon folding; a predicted bipartite NLS at the N-terminus of AID might not be functional (49) (Figure 4). In the nucleus, AID was recently found to accumulate in nucleolar structures where it associates with nucleolin and nucleophosmin (50). Mutations that abrogated AID localization to these structures resulted in reduced levels of CSR (50). The nucleoli may serve as a nucleation site for forming complexes, but the precise role of nucleolar AID remains unresolved.

A mutator protein's presence in the nucleus must be vigilantly regulated, and a nuclear export signal (NES) within the last 10 amino acids at the C-terminus of AID mediates CRM1-dependent active nuclear export (49) (Figure 4). Mutations in the NES increased levels of nuclear AID, enhanced SHM, but severely impaired CSR, indicating that NES-bearing C-terminus of AID plays a role in CSR beyond export, perhaps in mediating CSR-specific interactions (49). Consistent with this notion, replacement of the C-terminus of AID with a heterologous NES rescued nuclear export, but did not reconstitute CSR (51). Strikingly, the stability of AID was often compromised upon manipulation of the C-terminus of AID, even when nuclear export remained unaffected (51). In this regard, the half-life of nuclear AID is significantly shorter than its cytoplasmic counterpart (~2.5 vs. ~18 h) (52). This is largely due to interactions of AID with the proteasome through ubiquitination or Reg-γ-mediated escort (52, 53). Overall, the involvement of the C-terminus of AID in mediating nuclear export, protein stability, cytoplasmic retention, and CSR-specific interactions render this region one of the most fascinating, yet complicated domain that demands extensive examination.

#### AID PHOSPHORYLATION

Numerous putative phosphorylation sites in AID have been implicated in regulating its ability to effect CSR, SHM, and oncogenic translocations, without affecting stability or deamination potential



(**Figure 4**). Unfortunately, for most, mechanistic insights of functional pertinence remain elusive. Physiologically relevant sites that play critical to modest roles in AID function include serine-3 (S3), threonine-140 (T140), and S38 (54–58). Serine-3 was identified as a site phosphorylated by protein kinase C (PKC) *in vitro* (55). In contrast to other validated phosphorylation events, phosphorylated S3 inhibits AID function. Mutation of S3 to alanine enhances CSR, SHM, and *c-Myc/IgH* translocations, despite unperturbed catalytic activity (55); however, the mechanistic underpinnings remain unresolved. PKC can also phosphorylate T140, and T140A mutation perturbs SHM more profoundly than CSR. The mechanism through which phosphorylation at T140 differentially regulates SHM and CSR remains unclear (56). Phosphorylation of AID at serine-38 has been extensively characterized and will be discussed later. Overall, the regulatory mechanisms discussed above, and processes that mediate substrate specificity *in vivo* as discussed below, impose checkpoints in maintaining physiological functions of AID to facilitate successful and efficient CSR.

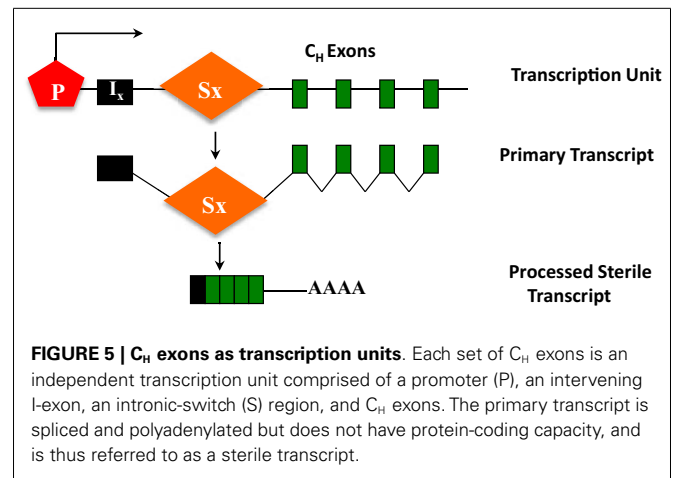
### ACCESSIBILITY AND TARGETING OF AID

Since AID is an ssDNA deaminase, mechanisms must exist to generate and reveal such structures at S regions during CSR. Additionally, AID must be actively and specifically recruited to S regions, not only to be productively engaged in CSR, but also to reduce collateral damage associated with expression of a mutator protein. The nature of S regions and their transcription promote AID accessibility to DNA while several proteins have been implicated to specifically recruit AID to the IgH locus during CSR.

### S REGIONS, TRANSCRIPTION, AND R-LOOPS

S regions are 1–12 kb repetitive sequences that are enriched with AID “hot-spot” 5′-RGYW-3′ motifs (59, 60), and are particularly G-rich on the non-template strand. Evidence for the role of S regions came from elegant genetic studies wherein deletion of S $\mu$  dramatically impaired CSR to all isotypes while deletion of S $\gamma$ 1 abolished CSR to IgG1 (61–63). Recent studies have provocatively suggested that apart from the default donor S $\mu$ , even S $\gamma$ 1 can serve as a donor and allow sequential switching to IgE (64), an idea that was suggested two decades back when double-isotype expressing B cells were identified (65–68).

The ability of S regions to serve as recombination targets is intricately linked with “germ-line” transcription, an essential prerequisite for CSR (2, 3). Each set of C<sub>H</sub> exons is an independent transcriptional unit, comprised of an intervening (I)-exon, intronic S region, and the C<sub>H</sub> exons (**Figure 5**). The primary transcripts produced constitutively (via  $\mu$  promoter) or inducibly (for other C<sub>H</sub> exons), are spliced and polyadenylated but have no protein-coding capacity. These are referred to as germ-line or sterile transcripts. Differential stimulation with distinct sets of activators and cytokines, provided by helper-T cells or through direct interaction with pathogens, induces transcription through different C<sub>H</sub> exons and promotes CSR to that particular isotype. Significant progress in our understanding of CSR came from *ex vivo* studies wherein splenic B cells were activated in culture under different conditions. For example, bacterial lipopolysaccharide (LPS) induces germ-line transcription through C $\gamma$ 2b and C $\gamma$ 3 and allows CSR to IgG2b and IgG3, while a combination



of LPS and interleukin-4 induces C $\gamma$ 1 and C $\epsilon$  transcription and CSR to IgG1 and IgE. Mutational analyses that altered or deleted the I-exon promoter perturbed CSR dramatically, thus providing experimental evidence for the strong mechanistic link between germ-line transcription and fulfillment of CSR (2, 3).

It is generally believed that transcription through S region sequences promotes formation of R-loops, wherein the template strand stably hybridizes with the G-rich primary transcript (69, 70). This allows the non-template strand to be looped out as ssDNA, providing an ideal substrate for AID (2) (**Figure 2**). Compelling work in support of the R-loop model came from the observations that a transcribed synthetic DNA fragment with a G-rich non-template strand can support AID deamination *in vitro* and CSR in B cells, while the inverted sequence (C-rich non-template strand) that does not form R-loops, neither supports AID-mediated deamination *in vitro* nor CSR *in vivo* (10, 63). It is to be noted that although the role for germ-line transcription has been well-studied, a possible role of the transcript *per se* was suggested from the observation that perturbing splicing of primary switch transcripts without affecting transcription impedes CSR (71, 72). However, the neomycin-resistance cassette used in targeting the splice donor site was not removed, leaving open the possibility that the observed CSR defect was due to non-specific effects of this cassette in the IgH locus. Despite this potential caveat, given how non-coding RNAs like HOTAIR and Xist drive PRC2 targeting (73, 74), it would not be surprising if these non-coding switch transcripts play a significant role in AID targeting and activity at S regions.

### FACTORS PROMOTING TEMPLATE STRAND DEAMINATION

The R-loop model does not account for the mechanism of template strand deamination by AID, a prerequisite for the formation of DSBs. Several models have been put forward to account for deamination of template strand. Anti-sense transcription through the IgH locus has been proposed to facilitate access of AID to the template strand (75); however, anti-sense transcription is not essential for CSR (76). Components of the RNA exosome complex have been shown to interact with AID and mediate accessibility to the template strand by degrading the nascent RNA hybridized to the template strand (77). Recent work cogently

elucidated that Nedd4-dependent ubiquitination modulates the fate of AID-associated RNA polymerase II (Pol II), thus generating free 3'-ends that serve as substrates for RNA exosomes (78). RNaseH has also been proposed to facilitate R-loop collapse to ensure template strand deamination (79). However, the kinetics of such R-loop degradation must be stringently regulated in the context of S regions to first allow AID to act on the non-template strand, and elucidation of such intricacies awaits future work.

### TARGETING AID TO DNA

The primary sequence of S regions, transcription, and R-loops set a platform favorable for AID activity. However, for AID to reach this platform inside the nucleus is analogous to finding a needle in a haystack. Although AID-instigated off-target breaks are incurred, the frequency is far less than what is observed for the Ig loci (80–82). The low abundance of AID at the non-Ig genes has led to the debate whether this represents true binding or mere background creeping into the chromatin immunoprecipitation-sequencing (ChIP-seq) analysis used in these experiments (83). While genome-wide occupancy studies suggested that AID associates with accessible chromatin at stalled promoters of transcribed genes (82), reanalysis of the same data set (83) contradicted the notion of genome-wide AID binding. The technicalities and subtleties of data normalization for ChIP-Seq studies seem to be at the heart of such disparate results, and do highlight the need for caution when interrogating chromatin binding of proteins with low nuclear abundance (83). Thus, while both genome-wide (82) and locus-specific ChIP (10, 81, 84) clearly show abundance of AID at the Ig loci, the efficiency of its binding to other genomic sequences needs to be re-evaluated. Nonetheless, AID-induced mutations at non-Ig genes are observed in even normal B cells (80). Thus, it is obvious that the process is stringently orchestrated to prevent bystander damage by AID. Several elements within the Ig loci have been implicated in targeting AID to the variable region exons during SHM (85); however, in this review we will primarily describe the factors that chaperone AID with exquisite precision to the S regions during CSR.

Activation-induced deaminase was shown to be in a complex with Pol II (84), and more specifically with Spt5, a Pol II-associated protein mechanistically linked to transcriptional pausing (81). Genome-wide Spt5 occupancy correlated significantly with stalled Pol II and was predictive of AID-dependent mutations. B cells depleted of Spt5 had a severe defect in CSR, a consequence of decreased AID binding to S regions (81). A comprehensive treatise on the role of RNA pol II pausing at S regions during CSR has been reviewed elsewhere (86). The germinal center-specific GANP protein has also been implicated in mediating AID–Spt5–Pol II interaction (87). However, GANP deficiency does not impair CSR. Thus, in the context of switching B cells, there might be other unidentified players that facilitate AID targeting to stalled Pol II, and recent studies have shown that members of the Pol II-associated factor 1 (PAF1) complex and histone chaperone FACT complex can promote immune diversification by regulating association of AID with Pol II (88).

The 14-3-3 adaptor proteins have been implicated in recruiting AID to DNA through their ability to interact with RGYW

sequences (89). It is unclear and somewhat counterintuitive as to what happens to 14-3-3 proteins after they chaperone AID to DNA, and why they do not compete directly with AID for DNA binding. Additionally, recent data suggests that 14-3-3 proteins perform scaffolding function by directly interacting with uracil DNA glycosylase (UNG) and protein kinase A (PKA), two proteins with well-established functions in CSR (90–92). Besides, the data implicate an AID C-terminus-dependent complex formation with 14-3-3 and subsequent targeting, but they fail to reconcile how an AID  $\Delta$ 189–198 mutant that is impaired in 14-3-3 binding, can be targeted to the S regions and generate mutations (93). Future work is warranted to unequivocally establish the role of the 14-3-3 adaptors in AID targeting.

Polypyrimidine-tract-binding protein-2 (PTBP2) was identified as an AID interactor that regulates AID targeting to S regions (94). Originally known to be a splicing regulator in brain (neuronal isoform, nPTB), this protein also interacted with both the sense and anti-sense S region transcripts in primary B cells undergoing CSR. Since splicing might be important for CSR (71), it is tempting to speculate a splicing regulation-associated function of PTBP2 in AID recruitment to S regions. Molecular insights into PTBP2-dependent regulation of AID targeting, and the fate of nuclear AID in the absence of PTBP2 will surely constitute the next phase of investigation.

### ROLE OF CHROMATIN MODIFICATIONS IN TARGETING AID

It is becoming increasingly clear that epigenetic marks play crucial roles in mediating S region accessibility (95, 96). Both donor and acceptor S regions are specifically enriched for acetylation and methylation marks at histones H3 and H4, generally associated with “open” chromatin, for example, H3K9/K14ac, H3K27ac, H4K8ac, and H3K4me3. It has been suggested that AID targeting to S $\mu$  is facilitated by the H3K9me3 mark, which tethers AID to the donor S region via the HP1–KAP1 complex (97). Additionally, PTIP, a component of the mixed-lineage leukemia-like complexes that are important regulators of H3K4 methylation, participates in CSR by regulating transcription-coupled chromatin accessibility. PTIP-deficient B cells have a severe defect in CSR due to decreased germ-line transcription of downstream C $H$  exons, and compromised DNA repair (98). Finally, combinatorial H3K9ac and H3S10 phosphorylation (H3K9acS10ph), specifically in the recombining S regions, deposited by GCN5/PCAF in stimulated B cells, leads to 14-3-3 adaptor-dependent AID binding to permit efficient CSR (96). However, it is to be noted that these chromatin marks are not likely to be unique to S regions, and thus cannot be sole determinants of regions permissive to AID activity. Interestingly, it has been shown that R-loops are tightly linked to H3S10ph, a chromatin condensation signature (99). Thus, it can be posited that R-loop formation facilitates H3S10ph chromatin modification, which in combination with H3K9me3 and H3K9ac marks, permits AID-mediated *in vivo* deamination of S region targets. The precise interplay of chromatin “writers,” “erasers,” and “readers” that regulate these events warrants further investigation, but it is unambiguous that this complex recombination reaction must be impeccably tuned by such epigenetic controls to prevent collateral damage by AID.



## GENERATION OF DOUBLE-STRAND BREAKS DOWNSTREAM OF DNA DEAMINATION

All the regulatory mechanisms alluded to above serve to generate AID-instigated dU lesions in S regions. Since CSR proceeds through DSB intermediates, the deaminated S regions need to be processed into DNA nicks, with two closely opposed nicks constituting a DSB (2, 3). This is achieved by components of the BER and MMR pathways.

### ROLE OF BER AND MMR PATHWAYS

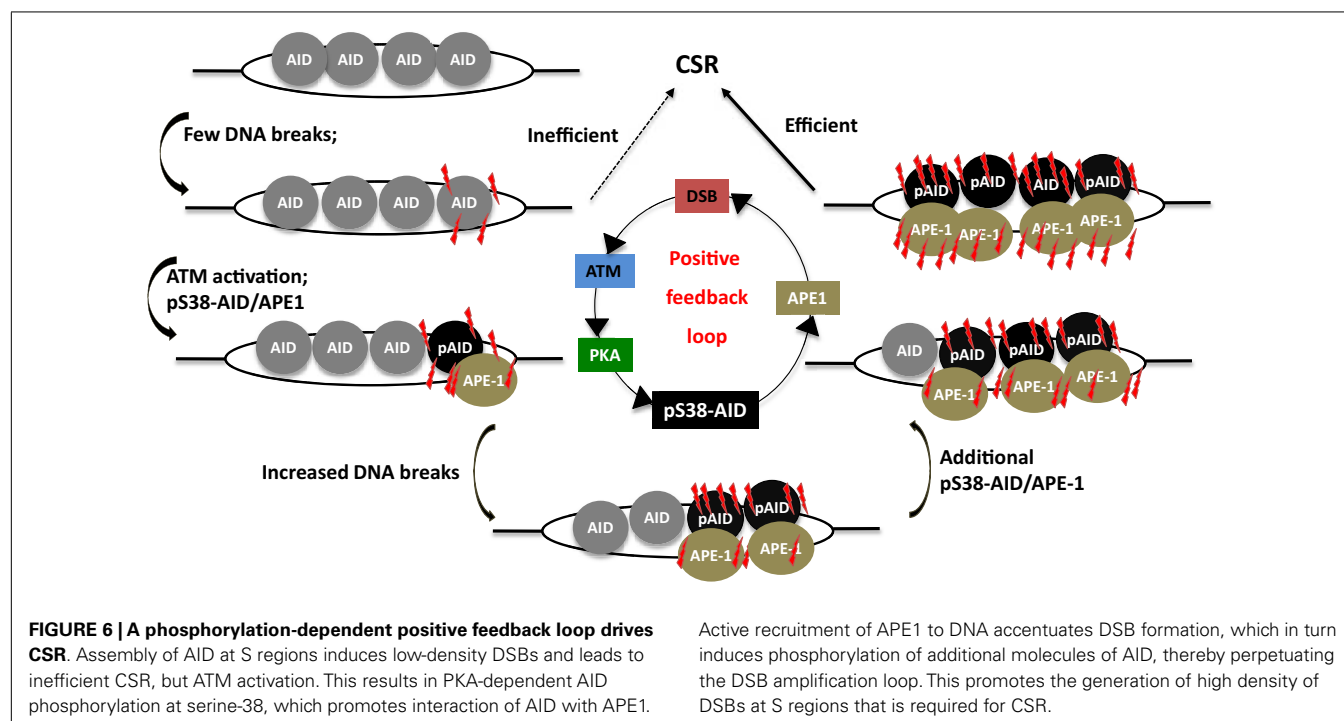
According to the prevailing model for CSR, UNG, a component of the BER pathway, removes the uracil base from deaminated S regions. The abasic site thus generated is converted into a nick by the apurinic/apyrimidinic endonuclease APE1. Two closely spaced nicks on opposite strands constitute a staggered DSB, further processing of which by nucleases or DNA polymerases (fill-in) generates a blunt DSB that can participate in end-joining (4, 5, 100). Consistent with this model, mutations in UNG lead to a severe defect in CSR, likely as a consequence of impaired formation of DSBs in S regions [reviewed in Ref. (2, 3)]. Additionally, APE1<sup>±</sup> mice and APE1-deficient CH12 cells reflected decreased DSBs in S regions and compromised CSR (101, 102). Components of MMR pathway have also been demonstrated to process DNA during CSR through the ability of Msh2:Msh6 to bind dU:dG mismatches, and subsequently recruit exonuclease 1 (exo1) to potentially process nicks and ssDNA gaps into DSBs. Indeed, mutations in Msh2 and Exo1 alter S region junctions and significantly impair CSR [reviewed in Ref. (2, 3)]. Conversely, deficiency of Pms2 and Mlh1, other members of the MMR machinery, lead to increased microhomology at S region junctions, suggesting that they might act to suppress alternative end-joining (103, 104). However, UNG mutations have a more profound effect on

CSR than mutations in MMR proteins, suggesting that CSR is more reliant on the UNG-dependent steps. Whether this reflects an uncharacterized preference for one pathway over the other or supports a proposed non-canonical role for UNG, independent of uracil removal activity during CSR, remains an open question (105).

### GENERATION OF HIGH DENSITY OF DSBs: REQUIREMENT OF AID PHOSPHORYLATION AT SERINE-38

The cellular DNA end-joining machinery is highly efficient and it is conceivable that a single DSB at an S region will be repaired before it can synapse with and ligate to a downstream DSB (106). It has therefore been speculated that efficient CSR would require a high density of DSBs at S regions to promote productive long distance synapsis and recombination between acceptor and donor S regions over intra-switch re-ligation (92), a phenomenon commonly observed in B cells that have initiated CSR but failed to complete the process (2). Recent studies have suggested that AID phosphorylated at serine-38 (S38) by PKA interacts with APE1 to actively generate a high density of breaks, a likely prerequisite for CSR (92, 107). In keeping with this notion, mutation of S38 to alanine severely impairs CSR due to a failure to efficiently generate DSBs at S regions (54, 56–58, 92, 107, 108).

Strikingly, AID phosphorylation at S38 was stimulated by DSBs (107). Thus, AID phosphorylation at S38 is both required for, and dependent on DSBs. This suggests the existence of a positive feedback loop wherein a low density of DSBs leads to AID phosphorylation, APE1 binding, and amplification of DSBs that feed-back into the loop (Figure 6). It was also demonstrated that ATM, a protein critical for cellular response to DNA-damage, participates in sensing the DSBs at S regions, thereby promoting AID phosphorylation and APE1 interaction (107). Being a



master regulator of the DNA-damage response, it is possible to envision ATM as a molecular rheostat that fine tunes DSB formation with efficient repair/recombination and allows safeguarded CSR while minimizing translocations. This is reminiscent of the role of RAG proteins in orchestrating V(D)J recombination by generating DSBs and efficiently channeling them to productive recombination, keeping translocation risks at bay (109–111). Similar to RAG-dependent coordination of break induction and repair, AID phosphorylated at S38 not only facilitates break formation, but also interacts with the ssDNA-binding protein, replication protein A, and likely enforces DNA repair pathways during CSR (112, 113).

Phosphorylation at S38 actively integrates AID functions into steps downstream of DNA deamination; however, several key questions remain elusive. First, the factors that facilitate APE1 binding to pS38AID need to be identified. Second, the regulatory mechanisms that couple chromatin sensing to DNA-damage signaling remain a mystery. Based on the recent finding of KAT5 (TIP60) tyrosine phosphorylation by DNA-damage to facilitate H3K9me3 binding and subsequent acetylation of ATM (114), it can be conjectured that such a pathway might be involved in the context of S regions and CSR, where the H3K9me3 mark has been shown to play a vital role (97). Finally, the steps between ATM activation and PKA-dependent AID phosphorylation remain a black box. These questions remain an active area of investigation.

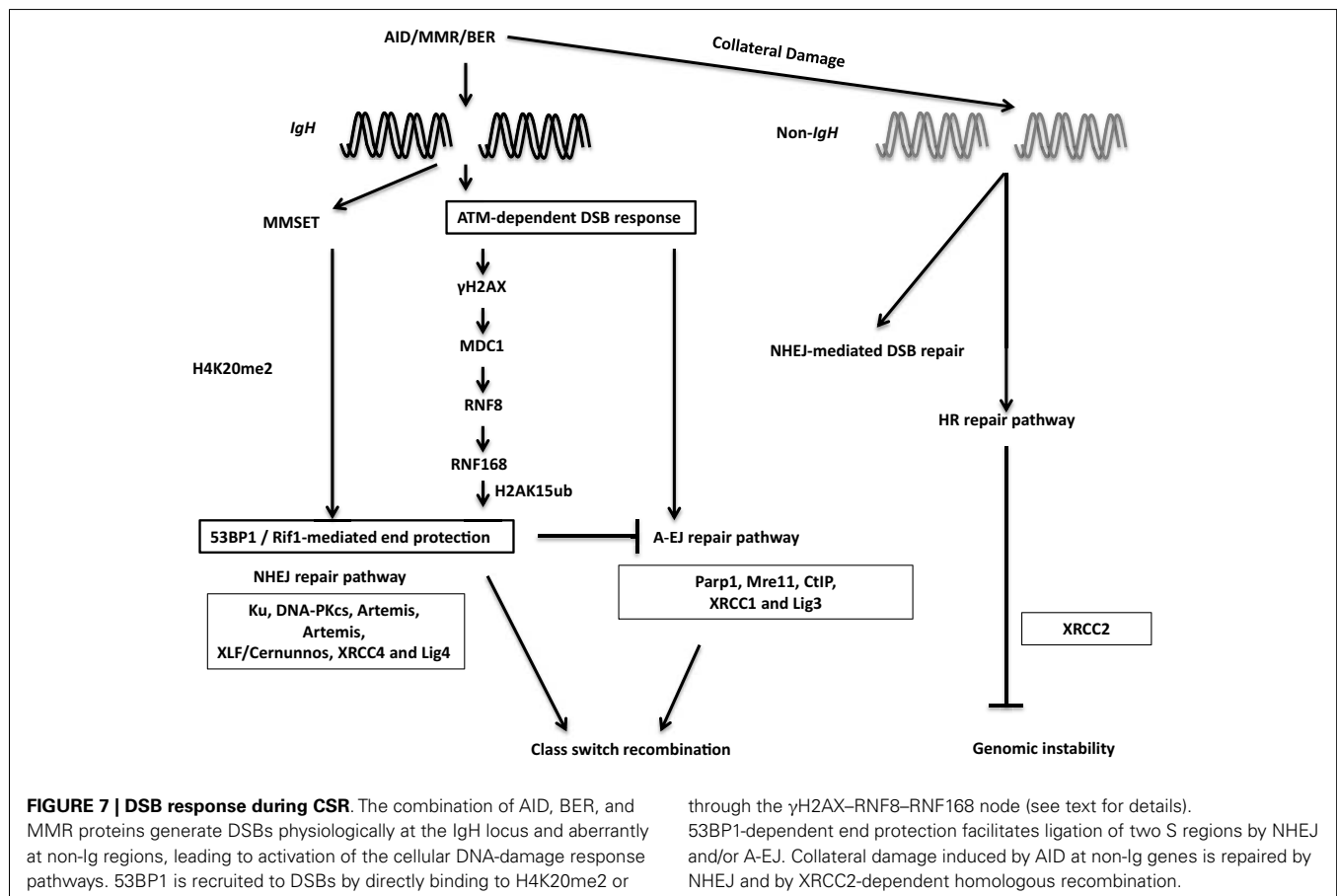
## COMPLETION OF CSR: END-JOINING OF SWITCH-REGION DSBs

Double-strand breaks generated at two distinct S regions are synapsed and ligated by end-joining during the completion phase of CSR (115). Below, we discuss the DSB response and DNA end-joining pathways that participate in this process.

### DSB RESPONSE DURING CSR

During the general DNA-damage response, DSBs are rapidly recognized by the Mre11–Rad50–Nbs1 (MRN) complex (116) (Figure 7). Nbs1 recruits and activates ATM, which phosphorylates H2AX. Phosphorylated H2AX ( $\gamma$ H2AX) serves as a docking site for several DNA response proteins and promotes the rapid accumulation of 53BP1, Nbs1, and MDC1 into repair foci near DSBs (116). Deletion or mutation of Nbs1, H2AX, 53BP1, and ATM impaired CSR, indicating that the proteins that participate in sensing and transducing DSBs participate in CSR (116). Additionally, the ATM-dependent DNA-damage response is required for maintenance of genomic integrity and suppression of oncogenic translocations, possibly through enforcing cell-cycle checkpoints (116). Overall, ATM promotes the assembly of macromolecular foci that stabilize DNA ends and facilitate the recruitment of repair factors to ensure productive CSR while preventing oncogenic translocations.

Among the ATM-activated DSB response factors, 53BP1 deficiency leads to the most pronounced defect in CSR (116). CSR



requires “synapsis” or close juxtaposition of donor and acceptor S regions (115, 117) and 53BP1 has been proposed to promote the synapsis of broken S regions during CSR (118). Furthermore, 53BP1 has been shown to associate with Rap1-interacting factor 1 (Rif1) to protect broken DNA ends from resection. Absence of Rif1 in B cells leads to increased DNA end resection, virtually phenocopying 53BP1 deficiency and providing functional significance to the 53BP1–Rif1 interaction during CSR (119–121). A recent study elegantly teased apart differential roles of distinct phospho-protein interactions of 53BP1, and convincingly illustrated that Rif1 serves as an effector of productive repair, whereas PTIP wards against mutagenic repair. This study clearly demonstrated that 53BP1 is a key player at the crossroads of efficient/aberrant DNA repair pathway choice (122).

The chromatin microenvironment strongly influences the DSB response. Two mechanisms have been proposed to regulate 53BP1 recruitment to DSBs in the context of chromatin. The first relies on its interaction with H4K20me2. Methylation of H4K20 and subsequent 53BP1 recruitment to sites of DNA-damage is regulated by MMSET, a histone methyltransferase (123, 124). MMSET depletion in the CH12F3 B cell line decreases H4K20me2 levels, attenuates 53BP1 accumulation at S regions, and impairs CSR (123). The other process that recruits 53BP1 to DSBs requires the RING finger protein 8 and 168 (RNF8 and RNF168)-dependent histone ubiquitination pathway. RNF8 is recruited to ATM-phosphorylated MDC1 bound to  $\gamma$ H2AX at the site of DSBs and catalyzes ubiquitin-dependent recruitment of RNF168 to chromatin flanking the DSBs (125). Recently, 53BP1 has been shown to recognize DNA-damage-induced H2A Lys15 ubiquitination catalyzed by RNF168, revealing the mechanism of RNF8/168-dependent recruitment of 53BP1 at DSBs. RNF8 deficiency compromised recruitment of 53BP1 to S regions in activated B cells and significantly impaired CSR. Additionally, inactivation of RNF168 impaired CSR in mice (126–129). Taken together, these observations suggest that 53BP1 recruitment plays a critical DSB end-protecting role during CSR.

### DNA END-JOINING

Non-homologous end-joining (NHEJ) is the primary mechanism used for end-joining during CSR (2). The canonical NHEJ machinery includes the Ku70/Ku80 heterodimer (Ku), DNA-dependent protein kinase catalytic subunit (DNA-PKcs), Artemis, XRCC4-like factor (XLF or Cernunnos), XRCC4, and DNA ligase IV (Lig4). Mutations in NHEJ components including Ku70/80, XRCC4, and DNA ligase IV severely compromise CSR. The role of NHEJ proteins in CSR is also evident from the observations that mutations in DNA-PKcs, artemis, and XLF lead to high levels of chromosomal IgH breaks and translocations, even in instances where CSR frequency is not severely impacted (116). Our current knowledge however, does not uncover how the initial recognition of DSBs by the MRN complex leads to the binding of Ku and DNA-PKcs to the broken DNA.

Non-homologous end-joining-deficiency does not abolish CSR, and S junctions in NHEJ-deficient B cells reveal extended microhomology, leading to the proposal that an alternative end-joining process (A-EJ) ligates DSBs during CSR (116). No factors unique to A-EJ have yet been characterized; several proteins

involved in other DNA repair pathways have been implicated to function in A-EJ, including Mre11 and CtIP. Mre11 and CtIP have been implicated to trim broken DNA ends to uncover microhomology regions, generating short stretches of complementary nucleotides at DNA breaks, thereby promoting A-EJ during CSR (116). In CH12F3 cells, CtIP depletion impaired CSR to IgA and reduced the overall length of microhomology at the S junctions (130, 131). Notably, CtIP-deficient B cells undergo normal CSR to IgG1 (132). Therefore, elucidation of the role of CtIP in CSR requires further investigation. A major open question relates to the interplay between NHEJ and A-EJ: does A-EJ operate in presence of intact NHEJ and does it have a physiological role other than being a mere backup to NHEJ?

### ROLE OF HOMOLOGY-DEPENDENT REPAIR IN RESOLVING AID-INDUCED BREAKS DURING CSR

It has been reported that AID instigates formation of widespread DSBs throughout the genome in activated B cells, albeit at significantly lower levels than that at the IgH locus (133–136). Such off-target DSBs at non-IgH loci are the major underlying lesions contributing to translocations between IgH and non-IgH loci (such as *c-Myc*) in B cells and are largely responsible for the ontogeny of a large number of B cell lymphomas in humans (21). In addition to aberrant translocations, AID can also induce somatic mutations at numerous loci linked to B cell tumorigenesis (80). While AID-initiated DSBs are observed in the G1 phase of the cell cycle (137), and CSR is likely completed before the cells transit into the S phase, it has been suggested that homologous recombination (HR)-dependent repair has a major role in providing resistance to AID-induced off-target DNA-damage. This is based on the observation that B cells deficient in the HR protein XRCC2 have significantly enhanced AID-dependent genome-wide DSBs (138) (Figure 7). Notably, the interplay between AID-mediated DNA breaks and HR repair pathway has been used in clinically relevant studies wherein AID-expressing human chronic lymphocytic leukemia cells were shown to be hypersensitive to HR inhibitors, possibly due to AID-dependent synthetic cytotoxicity (139). Further studies, in clinical settings, should be an interesting and possibly efficacious way to turn the mutator activity of AID into a therapy for B cell malignancies.

### PERSPECTIVE

The discovery of AID was a watershed event in the field of B cell biology and in deciphering the underlying cause behind the ontogeny of a large number of B cell lymphomas. We now have a working model of how non-coding transcription and DNA deamination initiate CSR and how general DNA repair proteins that function in distinct pathways contribute to the process. Still, a large number of unknowns plague the field. These include the mechanisms that specifically recruit AID to the Ig loci, leaving the rest of the genome largely untouched. We are yet to understand the processes that subvert normal DNA repair machineries and instead wield components of these pathways to promote recombination of DSBs that could be over 100 kb apart. The molecular basis underlying the balance between normal and aberrant repair requires further elucidation. Such basic knowledge can be exploited to shift the fulcrum of repair judiciously in clinical

settings of patients with immunodeficiency or lymphoid malignancies to reap translational benefit. Finally, DSB response occurs in, and is strongly influenced by the chromatin microenvironment. The dynamics of chromatin compaction and relaxation at DSBs are just beginning to unravel (140), but clearly much more remains to be unearthed as to how the dynamics of histone and DNA modifications impact and regulate a programmed DSB response that ensues during AID-orchestrated CSR. Addressing these exciting issues will be at the forefront of research in the coming years.

## ACKNOWLEDGMENTS

The authors wish to thank members of the Chaudhuri lab for helpful comments and discussion. This work was supported by grants from the National Institutes of Health (1R01AI072194) and the Starr Cancer Research Foundation to Jayanta Chaudhuri.

## REFERENCES

- Alt FW, Zhang Y, Meng FL, Guo C, Schwer B. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* (2013) **152**(3):417–29. doi:10.1016/j.cell.2013.01.007
- Chaudhuri J, Basu U, Zarrin A, Yan C, Franco S, Perlot T, et al. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv Immunol* (2007) **94**:157–214. doi:10.1016/S0065-2776(06)94006-1
- Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. *Annu Rev Immunol* (2008) **26**:261–92. doi:10.1146/annurev.immunol.26.021607.090248
- Maul RW, Saribasak H, Martomo SA, McClure RL, Yang W, Vaisman A, et al. Uracil residues dependent on the deaminase AID in immunoglobulin gene variable and switch regions. *Nat Immunol* (2011) **12**(1):70–6. doi:10.1038/ni.1970
- Petersen-Mahrt SK, Harris RS, Neuberger MS. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature* (2002) **418**(6893):99–103. doi:10.1038/nature00862
- Rada C, Di Noia JM, Neuberger MS. Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. *Mol Cell* (2004) **16**(2):163–71. doi:10.1016/j.molcel.2004.10.011
- Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem* (2007) **76**:1–22. doi:10.1146/annurev.biochem.76.061705.090740
- Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* (2000) **102**(5):553–63. doi:10.1016/S0092-8674(00)00078-7
- Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* (2000) **102**(5):565–75. doi:10.1016/S0092-8674(00)00079-9
- Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E, Alt FW. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* (2003) **422**(6933):726–30. doi:10.1038/nature01574
- Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the non-template strand. *Nat Immunol* (2003) **4**(5):452–6. doi:10.1038/ni920
- Kohli RM, Abrams SR, Gajula KS, Maul RW, Gearhart PJ, Stivers JT. A portable hot spot recognition loop transfers sequence preferences from APOBEC family members to activation-induced cytidine deaminase. *J Biol Chem* (2009) **284**(34):22898–904. doi:10.1074/jbc.M109.025536
- Vasudevan AA, Smits SH, Hoppner A, Haussinger D, Koenig BW, Munk C. Structural features of antiviral DNA cytidine deaminases. *Biol Chem* (2013) **394**(11):1357–70. doi:10.1515/hsz-2013-0165
- Pham P, Bransteitter R, Petruska J, Goodman MF. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* (2003) **424**(6944):103–7. doi:10.1038/nature01760
- Coker HA, Petersen-Mahrt SK. The nuclear DNA deaminase AID functions distributively whereas cytoplasmic APOBEC3G has a processive mode of action. *DNA Repair (Amst)* (2007) **6**(2):235–43. doi:10.1016/j.dnarep.2006.10.001
- Mak CH, Pham P, Afif SA, Goodman MF. A mathematical model for scanning and catalysis on single-stranded DNA, illustrated with activation-induced deoxycytidine deaminase. *J Biol Chem* (2013) **288**(41):29786–95. doi:10.1074/jbc.M113.506550
- Dickerson SK, Market E, Besmer E, Papavasiliou FN. AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med* (2003) **197**(10):1291–6. doi:10.1084/jem.20030481
- Kobayashi M, Aida M, Nagaoka H, Begum NA, Kitawaki Y, Nakata M, et al. AID-induced decrease in topoisomerase 1 induces DNA structural alteration and DNA cleavage for class switch recombination. *Proc Natl Acad Sci U S A* (2009) **106**(52):22375–80. doi:10.1073/pnas.0911879106
- Martin A, Scharff MD. Somatic hypermutation of the AID transgene in B and non-B cells. *Proc Natl Acad Sci U S A* (2002) **99**(19):12304–8. doi:10.1073/pnas.192442899
- Yoshikawa K, Okazaki IM, Eto T, Kinoshita K, Muramatsu M, Nagaoka H, et al. AID enzyme-induced hypermutation in an actively transcribed gene in fibroblasts. *Science* (2002) **296**(5575):2033–6. doi:10.1126/science.1071556
- Nussenzweig A, Nussenzweig MC. Origin of chromosomal translocations in lymphoid cancer. *Cell* (2010) **141**(1):27–38. doi:10.1016/j.cell.2010.03.016
- Lee-Theilen M, Chaudhuri J. Walking the AID tightrope. *Nat Immunol* (2010) **11**(2):107–9. doi:10.1038/ni2010-107
- Mai T, Zan H, Zhang J, Hawkins JS, Xu Z, Casali P. Estrogen receptors bind to and activate the HOXC4/HoxC4 promoter to potentiate HoxC4-mediated activation-induced cytosine deaminase induction, immunoglobulin class switch DNA recombination, and somatic hypermutation. *J Biol Chem* (2010) **285**(48):37797–810. doi:10.1074/jbc.M110.169086
- Pauklin S, Petersen-Mahrt SK. Progesterone inhibits activation-induced deaminase by binding to the promoter. *J Immunol* (2009) **183**(2):1238–44. doi:10.4049/jimmunol.0803915
- Pauklin S, Sernandez IV, Bachmann G, Ramiro AR, Petersen-Mahrt SK. Estrogen directly activates AID transcription and function. *J Exp Med* (2009) **206**(1):99–111. doi:10.1084/jem.20080521
- Huong le T, Kobayashi M, Nakata M, Shioi G, Miyachi H, Honjo T, et al. In vivo analysis of Aicda gene regulation: a critical balance between upstream enhancers and intronic silencers governs appropriate expression. *PLoS One* (2013) **8**(4):e61433. doi:10.1371/journal.pone.0061433
- Ise W, Kohyama M, Schraml BU, Zhang T, Schwer B, Basu U, et al. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. *Nat Immunol* (2011) **12**(6):536–43. doi:10.1038/ni.2037
- Betz BC, Jordan-Williams KL, Wang C, Kang SG, Liao J, Logan MR, et al. Batf coordinates multiple aspects of B and T cell function required for normal antibody responses. *J Exp Med* (2010) **207**(5):933–42. doi:10.1084/jem.20091548
- Dominguez-Sola D, Victoria GD, Ying CY, Phan RT, Saito M, Nussenzweig MC, et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat Immunol* (2012) **13**(11):1083–91. doi:10.1038/ni.2428
- Fernandez D, Ortiz M, Rodriguez L, Garcia A, Martinez D, Moreno de Alboran I. The proto-oncogene c-myc regulates antibody secretion and Ig class switch recombination. *J Immunol* (2013) **190**(12):6135–44. doi:10.4049/jimmunol.1300712
- Han JH, Akira S, Calame K, Beutler B, Selsing E, Imanishi-Kari T. Class switch recombination and somatic hypermutation in early mouse B cells are mediated by B cell and Toll-like receptors. *Immunity* (2007) **27**(1):64–75. doi:10.1016/j.immuni.2007.05.018
- Kumar S, Wuerffel R, Achour I, Lajoie B, Sen R, Dekker J, et al. Flexible ordering of antibody class switch and V(D)J joining during B-cell ontogeny. *Genes Dev* (2013) **27**(22):2439–44. doi:10.1101/gad.227165.113
- Endo Y, Marusawa H, Kinoshita K, Morisawa T, Sakurai T, Okazaki IM, et al. Expression of activation-induced cytidine deaminase in human hepatocytes via NF- $\kappa$ B signaling. *Oncogene* (2007) **26**(38):5587–95. doi:10.1038/sj.onc.1210344
- Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T, et al. *Helicobacter pylori* infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. *Nat Med* (2007) **13**(4):470–6. doi:10.1038/nm1566

35. Lin C, Yang L, Tanasa B, Hutt K, Ju BG, Ohgi K, et al. Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* (2009) **139**(6):1069–83. doi:10.1016/j.cell.2009.11.030
36. Rommel PC, Bosque D, Gitlin AD, Croft GF, Heintz N, Casellas R, et al. Fate mapping for activation-induced cytidine deaminase (AID) marks non-lymphoid cells during mouse development. *PLoS One* (2013) **8**(7):e69208. doi:10.1371/journal.pone.0069208
37. Kumar R, DiMenna L, Schrodde N, Liu T-C, Franck P, Munoz-Descalzo S, et al. AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature* (2013) **500**(7460):89–92. doi:10.1038/nature12299
38. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* (2004) **279**(50):52353–60. doi:10.1074/jbc.M407695200
39. Bhutani N, Brady JJ, Damian M, Sacco A, Corbel SY, Blau HM. Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* (2010) **463**(7284):1042–7. doi:10.1038/nature08752
40. Popp C, Dean W, Feng S, Cokus SJ, Andrews S, Pellegrini M, et al. Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* (2010) **463**(7284):1101–5. doi:10.1038/nature08829
41. Nabel CS, Jia H, Ye Y, Shen L, Goldschmidt HL, Stivers JT, et al. AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat Chem Biol* (2012) **8**(9):751–8. doi:10.1038/nchembio.1042
42. Fritz EL, Rosenberg BR, Lay K, Mihailovic A, Tuschl T, Papavasiliou FN. A comprehensive analysis of the effects of the deaminase AID on the transcriptome and methylome of activated B cells. *Nat Immunol* (2013) **14**(7):749–55. doi:10.1038/ni.2616
43. Dorsett Y, McBride KM, Jankovic M, Gazumyan A, Thai TH, Robbiani DF, et al. MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity* (2008) **28**(5):630–8. doi:10.1016/j.immuni.2008.04.002
44. Teng G, Hakimpour P, Landgraf P, Rice A, Tuschl T, Casellas R, et al. MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity* (2008) **28**(5):621–9. doi:10.1016/j.immuni.2008.03.015
45. Vigorito E, Perks KL, Abreu-Goodger C, Bunting S, Xiang Z, Kohlhaas S, et al. microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells. *Immunity* (2007) **27**(6):847–59. doi:10.1016/j.immuni.2007.10.009
46. Basso K, Schneider C, Shen Q, Holmes AB, Setty M, Leslie C, et al. BCL6 positively regulates AID and germinal center gene expression via repression of miR-155. *J Exp Med* (2012) **209**(13):2455–65. doi:10.1084/jem.20121387
47. de Yébenes VG, Belver L, Pisano DG, Gonzalez S, Villasante A, Croce C, et al. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J Exp Med* (2008) **205**(10):2199–206. doi:10.1084/jem.20080579
48. Loeb GB, Khan AA, Canner D, Hiatt JB, Shendure J, Darnell RB, et al. Transcriptome-wide miR-155 binding map reveals widespread noncanonical microRNA targeting. *Mol Cell* (2012) **48**(5):760–70. doi:10.1016/j.molcel.2012.10.002
49. Orthwein A, Di Noia JM. Activation induced deaminase: how much and where? *Semin Immunol* (2012) **24**(4):246–54. doi:10.1016/j.smim.2012.05.001
50. Hu Y, Ericsson I, Torseth K, Methot SP, Sundheim O, Liabakk NB, et al. A combined nuclear and nucleolar localization motif in activation-induced cytidine deaminase (AID) controls immunoglobulin class switching. *J Mol Biol* (2013) **425**(2):424–43. doi:10.1016/j.jmb.2012.11.026
51. Geisberger R, Rada C, Neuberger MS. The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proc Natl Acad Sci U S A* (2009) **106**(16):6736–41. doi:10.1073/pnas.0810808106
52. Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill JC, et al. Proteasomal degradation restricts the nuclear lifespan of AID. *J Exp Med* (2008) **205**(6):1357–68. doi:10.1084/jem.20070950
53. Uchimura Y, Barton LF, Rada C, Neuberger MS. REG-gamma associates with and modulates the abundance of nuclear activation-induced deaminase. *J Exp Med* (2011) **208**(12):2385–91. doi:10.1084/jem.20110856
54. Basu U, Chaudhuri J, Alpert C, Dutt S, Ranganath S, Li G, et al. The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature* (2005) **438**(7067):508–11. doi:10.1038/nature04255
55. Gazumyan A, Timachova K, Yuen G, Siden E, Di Virgilio M, Woo EM, et al. Amino-terminal phosphorylation of activation-induced cytidine deaminase suppresses c-myc/IgH translocation. *Mol Cell Biol* (2011) **31**(3):442–9. doi:10.1128/MCB.00349-10
56. McBride KM, Gazumyan A, Woo EM, Barreto VM, Robbiani DF, Chait BT, et al. Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. *Proc Natl Acad Sci U S A* (2006) **103**(23):8798–803. doi:10.1073/pnas.0603272103
57. McBride KM, Gazumyan A, Woo EM, Schwickert TA, Chait BT, Nussenzweig MC. Regulation of class switch recombination and somatic mutation by AID phosphorylation. *J Exp Med* (2008) **205**(11):2585–94. doi:10.1084/jem.20081319
58. Pasqualucci L, Kitanura Y, Gu H, Dalla-Favera R. PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. *Proc Natl Acad Sci U S A* (2006) **103**(2):395–400. doi:10.1073/pnas.0509969103
59. Rogozin IB, Diaz M. Cutting edge: DGYW/WRCY 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. *J Immunol* (2004) **172**(6):3382–4.
60. Zarrin AA, Alt FW, Chaudhuri J, Stokes N, Kaushal D, Du Pasquier L, et al. An evolutionarily conserved target motif for immunoglobulin class-switch recombination. *Nat Immunol* (2004) **5**(12):1275–81. doi:10.1038/ni1137
61. Khamlichi AA, Glaudet F, Oruc Z, Denis V, Le Bert M, Cogne M. Immunoglobulin class-switch recombination in mice devoid of any S mu tandem repeat. *Blood* (2004) **103**(10):3828–36. doi:10.1182/blood-2003-10-3470
62. Luby TM, Schrader CE, Stavnezer J, Selsing E. The mu switch region tandem repeats are important, but not required, for antibody class switch recombination. *J Exp Med* (2001) **193**(2):159–68. doi:10.1084/jem.193.2.159
63. 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**(5):435–41. doi:10.1038/ni918
64. Zhang T, Franklin A, Boboila C, McQuay A, Gallagher MP, Manis JP, et al. Downstream class switching leads to IgE antibody production by B lymphocytes lacking IgM switch regions. *Proc Natl Acad Sci U S A* (2010) **107**(7):3040–5. doi:10.1073/pnas.0915072107
65. Jabara HH, Loh R, Ramesh N, Vercelli D, Geha RS. Sequential switching from mu to epsilon via gamma 4 in human B cells stimulated with IL-4 and hydrocortisone. *J Immunol* (1993) **151**(9):4528–33.
66. Mills FC, Thyphronitis G, Finkelman FD, Max EE. Ig mu-epsilon isotype switch in IL-4-treated human B lymphoblastoid cells. Evidence for a sequential switch. *J Immunol* (1992) **149**(3):1075–85.
67. Yoshida K, Matsuoka M, Usuda S, Mori A, Ishizaka K, Sakano H. Immunoglobulin switch circular DNA in the mouse infected with *Nippostrongylus brasiliensis*: evidence for successive class switching from mu to epsilon via gamma 1. *Proc Natl Acad Sci U S A* (1990) **87**(20):7829–33. doi:10.1073/pnas.87.20.7829
68. Kunimoto DY, Sneller MC, Claflin L, Mushinski JF, Strober W. Molecular analysis of double isotype expression in IgA switching. *J Immunol* (1993) **150**(4):1338–47.
69. Tian M, Alt FW. Transcription-induced cleavage of immunoglobulin switch regions by nucleotide excision repair nucleases in vitro. *J Biol Chem* (2000) **275**(31):24163–72. doi:10.1074/jbc.M003343200
70. Yu K, Chedin F, Hsieh CL, Wilson TE, Lieber MR. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol* (2003) **4**(5):442–51. doi:10.1038/ni919
71. Lorenz M, Jung S, Radbruch A. Switch transcripts in immunoglobulin class switching. *Science* (1995) **267**(5205):1825–8. doi:10.1126/science.7892607
72. Hein K, Lorenz MG, Siebenkotten G, Petry K, Christine R, Radbruch A. Processing of switch transcripts is required for targeting of antibody class switch recombination. *J Exp Med* (1998) **188**(12):2369–74. doi:10.1084/jem.188.12.2369
73. Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* (2008) **322**(5902):750–6. doi:10.1126/science.1163045
74. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* (2010) **464**(7291):1071–6. doi:10.1038/nature08975
75. Perlot T, Li G, Alt FW. Antisense transcripts from immunoglobulin heavy-chain locus V(D)J and switch regions. *Proc Natl Acad Sci U S A* (2008) **105**(10):3843–8. doi:10.1073/pnas.0712291105
76. Haddad D, Dougier HL, Laviolette N, Puget N, Khamlichi AA. Replacement of Imu-Cmu intron by NeoR gene alters Imu germ-line expression but



- has no effect on V(D)J recombination. *Mol Immunol* (2010) **47**(5):961–71. doi:10.1016/j.molimm.2009.11.024
77. Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, et al. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* (2011) **144**(3):353–63. doi:10.1016/j.cell.2011.01.001
  78. Sun J, Keim CD, Wang J, Kazadi D, Oliver PM, Rabadan R, et al. E3-ubiquitin ligase Nedd4 determines the fate of AID-associated RNA polymerase II in B cells. *Genes Dev* (2013) **27**(16):1821–33. doi:10.1101/gad.210211.112
  79. Yu K, Lieber MR. Nucleic acid structures and enzymes in the immunoglobulin class switch recombination mechanism. *DNA Repair (Amst)* (2003) **2**(11):1163–74. doi:10.1016/j.dnarep.2003.08.010
  80. Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH, et al. Two levels of protection for the B cell genome during somatic hypermutation. *Nature* (2008) **451**(7180):841–5. doi:10.1038/nature06547
  81. Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, et al. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* (2010) **143**(1):122–33. doi:10.1016/j.cell.2010.09.017
  82. Yamane A, Resch W, Kuo N, Kuchen S, Li Z, Sun HW, et al. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat Immunol* (2011) **12**(1):62–9. doi:10.1038/ni.1964
  83. Hogenbirk MA, Velds A, Kerkhoven RM, Jacobs H. Reassessing genomic targeting of AID. *Nat Immunol* (2012) **13**(9):797–8; author reply 798–800. doi:10.1038/ni.2367
  84. Nambu Y, Sugai M, Gonda H, Lee CG, Katakai T, Agata Y, et al. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science* (2003) **302**(5653):2137–40. doi:10.1126/science.1092481
  85. Kothapalli NR, Fugmann SD. Targeting of AID-mediated sequence diversification to immunoglobulin genes. *Curr Opin Immunol* (2011) **23**(2):184–9. doi:10.1016/j.coi.2010.12.009
  86. Kenter AL. AID targeting is dependent on RNA polymerase II pausing. *Semin Immunol* (2012) **24**(4):281–6. doi:10.1016/j.smim.2012.06.001
  87. Singh SK, Maeda K, Eid MM, Almofty SA, Ono M, Pham P, et al. GANP regulates recruitment of AID to immunoglobulin variable regions by modulating transcription and nucleosome occupancy. *Nat Commun* (2013) **4**:1830. doi:10.1038/ncomms2823
  88. Willmann KL, Milosevic S, Pauklin S, Schmitz KM, Rangam G, Simon MT, et al. A role for the RNA pol II-associated PAF complex in AID-induced immune diversification. *J Exp Med* (2012) **209**(11):2099–111. doi:10.1084/jem.20112145
  89. Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, et al. 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. *Nat Struct Mol Biol* (2010) **17**(9):1124–35. doi:10.1038/nsmb.1884
  90. Lam T, Thomas LM, White CA, Li G, Pone EJ, Xu Z, et al. Scaffold functions of 14-3-3 adaptors in B cell immunoglobulin class switch DNA recombination. *PLoS One* (2013) **8**(11):e80414. doi:10.1371/journal.pone.0080414
  91. Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T, Neuberger MS. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr Biol* (2002) **12**(20):1748–55. doi:10.1016/S0960-9822(02)01215-0
  92. Vuong BQ, Lee M, Kabir S, Irimia C, Macchiarulo S, McKnight GS, et al. Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. *Nat Immunol* (2009) **10**(4):420–6. doi:10.1038/ni.1708
  93. Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM, Nussenzweig MC. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol Cell* (2003) **12**(2):501–8. doi:10.1016/S1097-2765(03)00309-5
  94. Nowak U, Matthews AJ, Zheng S, Chaudhuri J. The splicing regulator PTBP2 interacts with the cytidine deaminase AID and promotes binding of AID to switch-region DNA. *Nat Immunol* (2011) **12**(2):160–6. doi:10.1038/ni.1977
  95. Wang L, Wuerffel R, Feldman S, Khamlichi AA, Kenter AL. S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. *J Exp Med* (2009) **206**(8):1817–30. doi:10.1084/jem.20081678
  96. Li G, White CA, Lam T, Pone EJ, Tran DC, Hayama KL, et al. Combinatorial H3K9acS10ph histone modification in IgH locus S regions targets 14-3-3 adaptors and AID to specify antibody class-switch DNA recombination. *Cell Rep* (2013) **5**(3):702–14. doi:10.1016/j.celrep.2013.09.031
  97. Jeevan-Raj BP, Robert I, Heyer V, Page A, Wang JH, Cammas F, et al. Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *J Exp Med* (2011) **208**(8):1649–60. doi:10.1084/jem.20110118
  98. Daniel JA, Santos MA, Wang Z, Zang C, Schwab KR, Jankovic M, et al. PTIP promotes chromatin changes critical for immunoglobulin class switch recombination. *Science* (2010) **329**(5994):917–23. doi:10.1126/science.1187942
  99. Castellano-Pozo M, Santos-Pereira JM, Rondon AG, Barroso S, Andujar E, Perez-Alegre M, et al. R loops are linked to histone h3 s10 phosphorylation and chromatin condensation. *Mol Cell* (2013) **52**(4):583–90. doi:10.1016/j.molcel.2013.10.006
  100. Chaudhuri J, Alt FW. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat Rev Immunol* (2004) **4**(7):541–52. doi:10.1038/nri1425
  101. Masani S, Han L, Yu K. Apurinic/aprymidinic endonuclease 1 is the essential nuclease during immunoglobulin class switch recombination. *Mol Cell Biol* (2013) **33**(7):1468–73. doi:10.1128/MCB.00026-13
  102. Schrader CE, Guikema JE, Wu X, Stavnezer J. The roles of APE1, APE2, DNA polymerase beta and mismatch repair in creating S region DNA breaks during antibody class switch. *Philos Trans R Soc Lond B Biol Sci* (2009) **364**(1517):645–52. doi:10.1098/rstb.2008.0200
  103. Ehrenstein MR, Rada C, Jones AM, Milstein C, Neuberger MS. Switch junction sequences in PMS2-deficient mice reveal a microhomology-mediated mechanism of Ig class switch recombination. *Proc Natl Acad Sci U S A* (2001) **98**(25):14553–8. doi:10.1073/pnas.241525998
  104. Schrader CE, Vardo J, Stavnezer J. Role for mismatch repair proteins Msh2, Mlh1, and Pms2 in immunoglobulin class switching shown by sequence analysis of recombination junctions. *J Exp Med* (2002) **195**(3):367–73. doi:10.1084/jem.20011877
  105. Begum NA, Stanlie A, Doi T, Sasaki Y, Jin HW, Kim YS, et al. Further evidence for involvement of a noncanonical function of uracil DNA glycosylase in class switch recombination. *Proc Natl Acad Sci U S A* (2009) **106**(8):2752–7. doi:10.1073/pnas.0813252106
  106. Rush JS, Fugmann SD, Schatz DG. Staggered AID-dependent DNA double strand breaks are the predominant DNA lesions targeted to Sm in immunoglobulin class switch recombination. *Int Immunol* (2004) **16**:549–57. doi:10.1093/intimm/dxh057
  107. Vuong BQ, Herrick-Reynolds K, Vaidyanathan B, Pucella JN, Ucher AJ, Donghia NM, et al. A DNA break- and phosphorylation-dependent positive feedback loop promotes immunoglobulin class-switch recombination. *Nat Immunol* (2013) **14**(11):1183–9. doi:10.1038/ni.2732
  108. Cheng HL, Vuong BQ, Basu U, Franklin A, Schwer B, Astarita J, et al. Integrity of the AID serine-38 phosphorylation site is critical for class switch recombination and somatic hypermutation in mice. *Proc Natl Acad Sci U S A* (2009) **106**(8):2717–22. doi:10.1073/pnas.0812304106
  109. Lee GS, Neiditch MB, Salus SS, Roth DB. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* (2004) **117**(2):171–84. doi:10.1016/S0092-8674(04)00301-0
  110. Chaumeil J, Micsinai M, Ntziachristos P, Roth DB, Aifantis I, Kluger Y, et al. The RAG2 C-terminus and ATM protect genome integrity by controlling antigen receptor gene cleavage. *Nat Commun* (2013) **4**:2231. doi:10.1038/ncomms3231
  111. Deriano L, Chaumeil J, Coussens M, Multani A, Chou Y, Alekseyenko AV, et al. The RAG2 C terminus suppresses genomic instability and lymphomagenesis. *Nature* (2011) **471**(7336):119–23. doi:10.1038/nature09755
  112. Chaudhuri J, Khuong C, Alt FW. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature* (2004) **430**(7003):992–8. doi:10.1038/nature02821
  113. Yamane A, Robbiani DF, Resch W, Bothmer A, Nakahashi H, Oliveira T, et al. RPA accumulation during class switch recombination represents 5'-3' DNA-end resection during the S-G2/M phase of the cell cycle. *Cell Rep* (2013) **3**(1):138–47. doi:10.1016/j.celrep.2012.12.006
  114. Kaidi A, Jackson SP. KAT5 tyrosine phosphorylation couples chromatin sensing to ATM signalling. *Nature* (2013) **498**(7452):70–4. doi:10.1038/nature12201
  115. Wuerffel R, Wang L, Grigera F, Manis J, Selsing E, Perlot T, et al. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity* (2007) **27**(5):711–22. doi:10.1016/j.immuni.2007.09.007

116. Boboila C, Alt FW, Schwer B. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv Immunol* (2012) **116**:1–49. doi:10.1016/B978-0-12-394300-2.00001-6
117. Sellars M, Reina-San-Martin B, Kastner P, Chan S. Ikaros controls isotype selection during immunoglobulin class switch recombination. *J Exp Med* (2009) **206**(5):1073–87. doi:10.1084/jem.20082311
118. Reina-San-Martin B, Difilippantonio S, Hanitsch L, Masilamani RF, Nussenzweig A, Nussenzweig MC. H2AX is required for recombination between immunoglobulin switch regions but not for intra-switch region recombination or somatic hypermutation. *J Exp Med* (2003) **197**(12):1767–78. doi:10.1084/jem.20030569
119. Escribano-Diaz C, Orthwein A, Fradet-Turcotte A, Xing M, Young JT, Tkac J, et al. A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Mol Cell* (2013) **49**(5):872–83. doi:10.1016/j.molcel.2013.01.001
120. Chapman JR, Barral P, Vannier JB, Borel V, Steger M, Tomas-Loba A, et al. RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. *Mol Cell* (2013) **49**(5):858–71. doi:10.1016/j.molcel.2013.01.002
121. Zimmermann M, Lottersberger F, Buonomo SB, Sfeir A, de Lange T. 53BP1 regulates DSB repair using Rif1 to control 5' end resection. *Science* (2013) **339**(6120):700–4. doi:10.1126/science.1231573
122. Callen E, Di Virgilio M, Kruhlak MJ, Nieto-Soler M, Wong N, Chen HT, et al. 53BP1 mediates productive and mutagenic DNA repair through distinct phosphoprotein interactions. *Cell* (2013) **153**(6):1266–80. doi:10.1016/j.cell.2013.05.023
123. Pei H, Zhang L, Luo K, Qin Y, Chesi M, Fei F, et al. MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites. *Nature* (2011) **470**(7332):124–8. doi:10.1038/nature09658
124. Hajdu I, Ciccio A, Lewis SM, Elledge SJ. Wolf-Hirschhorn syndrome candidate 1 is involved in the cellular response to DNA damage. *Proc Natl Acad Sci U S A* (2011) **108**(32):13130–4. doi:10.1073/pnas.1110081108
125. Stewart GS. Solving the RIDDLE of 53BP1 recruitment to sites of damage. *Cell Cycle* (2009) **8**(10):1532–8. doi:10.4161/cc.8.10.8351
126. Li L, Halaby MJ, Hakem A, Cardoso R, El Ghamrasni S, Harding S, et al. Rnf8 deficiency impairs class switch recombination, spermatogenesis, and genomic integrity and predisposes for cancer. *J Exp Med* (2010) **207**(5):983–97. doi:10.1084/jem.20092437
127. Ramachandran S, Chahwan R, Nepal RM, Frieder D, Panier S, Roa S, et al. The RNF8/RNF168 ubiquitin ligase cascade facilitates class switch recombination. *Proc Natl Acad Sci U S A* (2010) **107**(2):809–14. doi:10.1073/pnas.0913790107
128. Santos MA, Huen MS, Jankovic M, Chen HT, Lopez-Contreras AJ, Klein IA, et al. Class switching and meiotic defects in mice lacking the E3 ubiquitin ligase RNF8. *J Exp Med* (2010) **207**(5):973–81. doi:10.1084/jem.20092308
129. Bohgaki T, Bohgaki M, Cardoso R, Panier S, Zeegers D, Li L, et al. Genomic instability, defective spermatogenesis, immunodeficiency, and cancer in a mouse model of the RIDDLE syndrome. *PLoS Genet* (2011) **7**(4):e1001381. doi:10.1371/journal.pgen.1001381
130. Lee-Theilen M, Matthews AJ, Kelly D, Zheng S, Chaudhuri J. CtIP promotes microhomology-mediated alternative end joining during class-switch recombination. *Nat Struct Mol Biol* (2011) **18**(1):75–9. doi:10.1038/nsmb.1942
131. Cortizas EM, Zahn A, Hajjar ME, Patenaude AM, Di Noia JM, Verdun RE. Alternative end-joining and classical nonhomologous end-joining pathways repair different types of double-strand breaks during class-switch recombination. *J Immunol* (2013) **191**(11):5751–63. doi:10.4049/jimmunol.1301300
132. Bothmer A, Rommel PC, Gazumyan A, Polato F, Reczek CR, Muellenbeck MF, et al. Mechanism of DNA resection during intrachromosomal recombination and immunoglobulin class switching. *J Exp Med* (2013) **210**(1):115–23. doi:10.1084/jem.20121975
133. Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, et al. Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* (2011) **147**(1):107–19. doi:10.1016/j.cell.2011.07.049
134. Hasham MG, Donghia NM, Coffey E, Maynard J, Snow KJ, Ames J, et al. Widespread genomic breaks generated by activation-induced cytidine deaminase are prevented by homologous recombination. *Nat Immunol* (2010) **11**(9):820–6. doi:10.1038/ni.1909
135. Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, et al. Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell* (2011) **147**(1):95–106. doi:10.1016/j.cell.2011.07.048
136. Staszewski O, Baker RE, Ucher AJ, Martier R, Stavnezer J, Guikema JE. Activation-induced cytidine deaminase induces reproducible DNA breaks at many non-Ig loci in activated B cells. *Mol Cell* (2011) **41**(2):232–42. doi:10.1016/j.molcel.2011.01.007
137. Petersen S, Casellas R, Reina-San-Martin B, Chen HT, Difilippantonio MJ, Wilson PC, et al. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature* (2001) **414**(6864):660–5. doi:10.1038/414660a
138. Hasham MG, Snow KJ, Donghia NM, Branca JA, Lessard MD, Stavnezer J, et al. Activation-induced cytidine deaminase-initiated off-target DNA breaks are detected and resolved during S phase. *J Immunol* (2012) **189**(5):2374–82. doi:10.4049/jimmunol.1200414
139. Lamont KR, Hasham MG, Donghia NM, Branca J, Chavaree M, Chase B, et al. Attenuating homologous recombination stimulates an AID-induced antileukemic effect. *J Exp Med* (2013) **210**(5):1021–33. doi:10.1084/jem.20121258
140. Price BD, D'Andrea AD. Chromatin remodeling at DNA double-strand breaks. *Cell* (2013) **152**(6):1344–54. doi:10.1016/j.cell.2013.02.011

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 January 2014; accepted: 07 March 2014; published online: 28 March 2014.  
Citation: Vaidyanathan B, Yen W-F, Pucella JN and Chaudhuri J (2014) AIDing chromatin and transcription-coupled orchestration of immunoglobulin class-switch recombination. *Front. Immunol.* 5:120. doi: 10.3389/fimmu.2014.00120

This article was submitted to *B Cell Biology*, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Vaidyanathan, Yen, Pucella and Chaudhuri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Epigenetic regulation of individual modules of the immunoglobulin heavy chain locus 3' regulatory region

Barbara K. Birshtein\*

Department of Cell Biology, Albert Einstein College of Medicine, Bronx, NY, USA

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

James Hagman, National Jewish Health, USA

Paolo Casali, University of Texas Health Science Center, USA

## \*Correspondence:

Barbara K. Birshtein, Department of Cell Biology, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA  
e-mail: barbara.birshtein@einstein.yu.edu

The *Igh* locus undergoes an amazing array of DNA rearrangements and modifications during B cell development. During early stages, the variable region gene is constructed from constituent *variable* (*V*), *diversity* (*D*), and *joining* (*J*) segments (*VDJ* joining). B cells that successfully express an antibody can be activated, leading to somatic hypermutation (SHM) focused on the variable region, and class switch recombination (CSR), which substitutes downstream constant region genes for the originally used *C $\mu$*  constant region gene. Many investigators, ourselves included, have sought to understand how these processes specifically target the *Igh* locus and avoid other loci and potential deleterious consequences of malignant transformation. Our laboratory has concentrated on a complex regulatory region (RR) that is located downstream of *C $\alpha$* , the most 3' of the *Igh* constant region genes. The ~40 kb 3' RR, which is predicted to serve as a downstream major regulator of the *Igh* locus, contains two distinct segments: an ~28 kb region comprising four enhancers, and an adjacent ~12 kb region containing multiple CTCF and Pax5 binding sites. Analysis of targeted mutations in mice by a number of investigators has concluded that the entire 3' RR enhancer region is essential for SHM and CSR (but not for *VDJ* joining) and for high levels of expression of multiple isotypes. The CTCF/Pax5 binding region is a candidate for influencing *VDJ* joining early in B cell development and serving as a potential insulator of the *Igh* locus. Components of the 3' RR are subject to a variety of epigenetic changes during B cell development, i.e., DNase I hypersensitivity, histone modifications, and DNA methylation, in association with transcription factor binding. I propose that these changes provide a foundation by which regulatory elements in modules of the 3' RR function by interacting with each other and with target sequences of the *Igh* locus.

**Keywords:** immunoglobulin heavy chain gene locus, enhancers, insulators, CTCF, Pax5, class switch recombination, somatic hypermutation

## DISCOVERY OF 3' RR ENHANCERS

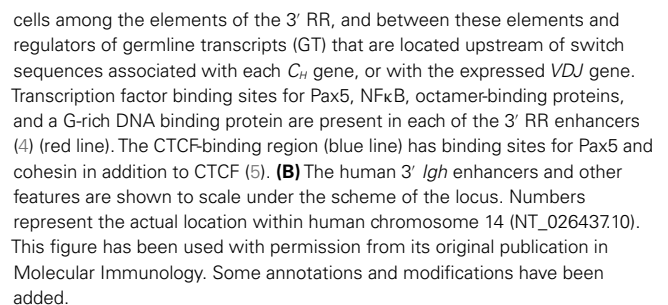
The *Igh* locus spans ~3 Mb, beginning near the telomere on murine chromosome 12 with the component *variable* (*V*), *diversity* (*D*), and *joining* (*J*) segments of the variable region, followed by the multiple constant region (*C<sub>H</sub>*) genes (Figure 1). My laboratory has been interested in the regulation of the *Igh* locus's multiple recombination and mutation processes that generate a diverse antigen-recognition repertoire. The entire 3' *Igh* regulatory region (RR) (enhancers and insulators) has been shown by others to potentially contribute to regulation of variable region formation (*VDJ* joining) (1). Importantly, it is definitively essential for class switch recombination (CSR) (2), and somatic hypermutation (SHM) (3). This review focuses on our studies on the structure and epigenetic regulation of the 3' RR as it contributes to those antibody diversification processes.

The first transcriptional enhancer identified in mammalian cells was the intronic enhancer of the *Igh* locus (*E $\mu$* ), positioned between the 3'-most *J<sub>H</sub>* segment and the 5'-most *C<sub>H</sub>* region, *C $\mu$*  [reviewed in Ref. (6)]. *E $\mu$*  was found to confer expression upon *Igh* genes when transfected into B cells, and was generally considered

to be of critical importance in enabling B cell-specific expression of the *Igh* locus.

Not surprisingly, perhaps, *E $\mu$*  was not the only B cell-specific enhancer in the *Igh* locus. When B cell lines that had deleted *E $\mu$*  were found to retain the ability to express the *Igh* gene (7, 8), the questions of gene regulation of *Igh* genes became increasingly provocative. What exactly was *E $\mu$* 's role? Was *E $\mu$*  required to initiate *Igh* expression but not to maintain it? Were there additional enhancers that compensated for the absence of *E $\mu$* , and where in the *Igh* locus might they be found? Examining a rat cosmid, the Neuberger group identified a DNA sequence with B cell-specific enhancer activity that was located ~25 kb downstream of *C $\alpha$* , the most 3' of the *C<sub>H</sub>* genes (9): this was the first of the 3' enhancers to be identified.

This newly identified enhancer was satisfyingly predicted to account for the aberrant expression of *myc* in various B cell malignancies when *myc* was activated as an oncogene via chromosomal translocation with the *Igh* locus. The translocation breakpoints in switch sequences upstream of *C<sub>H</sub>* genes divorced the intronic enhancer from the oncogenic transcription unit, leaving *myc*



by various investigators, using genomic sequences that were only then becoming identified, additional DNase I hypersensitive sites 3' of *C $\alpha$*  were detected, primarily using the mouse locus as a model (12–15). A mouse BAC sequence identified by Roy Riblet was found to encompass the entire 3' RR and the nearest downstream non-*Igh* genes (AF450245) (16). Similar experiments identified analogous enhancers of the human *Igh* locus (17–19), with the enhancer-containing segment of the human 3' RR fully characterized by the Max laboratory (20). **Figure 1** shows the general features of the enhancer-containing segments of the 3' RR in both mouse and human. A CTCF/Pax5 binding region

with insulator activity is located within the 3' RR downstream of the 3' RR enhancers (hs5–8), and will be discussed further below.

### STRUCTURAL FEATURES OF THE 3' RR ENHANCER-CONTAINING REGIONS

There are noteworthy structural features of the 3' RR (note that this region is sometimes referred to as 3' E $\alpha$ ) [reviewed in Ref. (21–24) (**Figure 1**)]. (1) Multiple DNase I HS sites with enhancer activity are dispersed in relatively large DNA segments (~28 kb in mouse) – a total of four enhancers in mouse (hs3a, hs1.2, hs3b, and hs4) and three in humans (hs1.2, hs3, and hs4). (2) In humans, there are two individual 3' RRs, one each downstream of *Ca1* and *Ca2*, respectively. They are quite similar to each other in sequence. The orientation of hs1.2 with respect to upstream and downstream sequences is reversed in the two 3' RRs in human, and also between rat and mouse 3' RRs (18). (3) A conserved palindrome feature, although not its specific sequences, flanks the central enhancer – hs1.2 (25). In mouse, the palindrome extends in both directions from hs1.2 to terminate at two virtually identical enhancers, hs3a and hs3b (26). Compared to mouse, the hs1.2 palindromic region in humans is shorter (27). Other species also have a conserved palindrome (28). (4) The hs4 enhancer is located outside and downstream of the palindrome. (5) Individual 3' RR enhancers in a given species, like hs1.2, hs3, and hs4, differ in sequence from each other (except for the virtually identical hs3a and hs3b in rodents). (6) There are limited homologies in enhancer sequence between species (e.g., hs1.2 in human and hs1.2 in mouse) (27).

Other than revealing a conserved palindromic structure, the regions between the 3' RR enhancers show virtually no homology between rodents and humans (27). Nonetheless, other particular sequence features stand out, as identified through genomic Southern analysis and percentage identity (dot-plot) analysis (29). In mouse (and rat), the “palindromic” sequences that separate hs1.2 from each of the terminal enhancers at the end of the palindrome, hs3a and hs3b, contain families of direct and inverted repeats (26) while the human (and chimpanzee) 3'RR revealed several regions of repetitive switch-like sequences (27). More recently, the Cogne laboratory specifically sought and identified multiple switch-gamma 1-like repeats in the mouse and human 3' RRs that were situated close to each of the four enhancers, as well as less distinct although evident, similar sequences in other species, like rabbit and dog (30). In mouse, the 3' RR is highly polymorphic (26, 31), showing variations in the lengths of the sequences between the enhancers and the number of repeats in these regions. The hs1.2 region in humans is polymorphic, with varying frequency of alleles in different populations (32). Polymorphic patterns of human hs1.2, i.e., alleles, are associated with different autoimmune disorders, such as lupus (33).

In summary, the 3' RR contains several enhancers located in two structurally distinctive modules – (1) a palindromic region (mouse hs3a–hs1.2–hs3b) and (2) a separate structural unit (hs4). Inter-enhancer regions reveal repetitive, switch-like sequences potentially of functional significance for the *Igh* locus. Downstream of the enhancer-containing segment of the 3' RR is additional DNase I hypersensitive sites (hs5–8), which contain CTCF and Pax5 sites

and have insulator activity. This hs5–8 region is discussed more fully later.

### TRANSCRIPTIONAL REGULATION OF 3' RR ENHANCERS

Relatively coincident with these studies on the *Igh* locus were studies of genes of the  $\beta$ -globin locus, which, like the *Igh* genes, are subject to developmental regulation. Multiple DNase I hypersensitive sites, each with enhancer activity, are located upstream of the  $\beta$ -globin genes. This enhancer-containing region is referred to as the locus control region because endogenous deletions here are found to affect expression of distally situated globin genes. Experimental questions on *Igh* genes have paralleled experiments carried out in the  $\beta$ -globin locus and in other loci [recent review in Ref. (34)]: (1) what are the protein factors that bind to and regulate these *Igh* enhancers? Can they account for B cell-specific regulation? (2) Are the different enhancers similar in their function, their relative “strength” and their activity on the target *Igh* locus? How do these enhancers work together? Can these questions be answered not only for *in vitro*, cellular conditions but also within the animal context itself?

Electrophoretic mobility shift assays (EMSA) provided a tool to identify proteins with the potential to bind enhancers. Using nuclear extracts, we identified a B cell-specific binding protein with sites throughout the *Igh* locus, including hs1.2 (35, 36). Based on its cellular expression pattern, we predicted that this protein was B cell-specific activating protein (BSAP), now called Pax5, as originally identified by the Busslinger laboratory (37); and various observations were consistent with that prediction (36). Additional 3' RR binding factors were identified leading to recognition of a quartet of proteins – Pax5, octamer-binding proteins, NF $\kappa$ B, and a G-rich DNA binding protein – that worked together on each of the murine 3' enhancers (4) (**Figure 1**).

Our experiments revealed that BSAP/Pax5 bound to each of the mouse 3' RR enhancers, where it could act as a repressor or activator (4). For example, mutational inactivation of a BSAP/Pax5 binding site of hs1.2 resulted in an *increase* in hs1.2 enhancer activity upon transfection into B cell lines that expressed endogenous BSAP/Pax5 (36). This finding showed that “BSAP” could be a repressor of hs1.2. The enhancer activity resulting from mutation of the BSAP/Pax5 binding site depended on the binding of the remaining transcription factors (38). A similar outcome applied not only to BSAP/Pax5 but also to each individual component of this quartet, as individual mutation of other binding sites each resulted in an increase in hs1.2 enhancer activity (38). Collectively, then, this quartet worked in concert to repress hs1.2, while it activated hs4, revealing that individual 3' RR enhancers had different B cell-specific activities (4). Interestingly, human 3' enhancers do not have Pax5 binding sites, suggesting that humans and mice have different modes of 3' RR regulation. However, it is not known how or whether the differences in Pax5 binding affect the function of the 3' RR in human and mouse. Human hs4 showed binding to octamer-binding proteins, NF $\kappa$ B, and YY1 under some circumstances, and human hs1.2 to octamer-binding proteins and Spi1, and to NF $\kappa$ B for some of the polymorphic hs1.2 variants (33, 39–41). Other 3' enhancer binding proteins have also been identified (40, 42).



## ADDITIONAL MODULE DOWNSTREAM OF 3' RR ENHANCERS: THE CTCF-BINDING REGION OF THE 3' RR

Had we identified all the regulators of the 3' RR [reviewed in Ref. (43)]? Various observations suggested that additional functional motifs were present beyond hs4. For example, the nearest non-*Igh* genes downstream of hs4, i.e., *hole* (*Tmem121*), *Crip*, and *mta-1*, each had a transcriptional orientation that was opposite to that of all of the *V*, *D*, *J*, and *C<sub>H</sub>* elements of the *Igh* locus. This back-to-back orientation led us to predict that a terminus of the *Igh* locus might be located in this segment (16). In fact, we found additional DNase I hypersensitive sites downstream of hs4, which included hs5, 6, and 7, and has now been extended to include a CTCF-binding site, named hs8 (44, 45). Discussions via Sandy Morse with Victor Lobanenkov introduced us to CTCF as a mammalian insulator (46), and we predicted that CTCF sites might be present in this region. EMSA of 50 overlapping DNA sequences with recombinant CTCF revealed a CTCF-binding module of the 3' RR [recently referred to as 3' CBE, CTCF-binding elements (47)], and transient transfection assays confirmed functional insulator activity in the absence of any enhancer activity (44). The CTCF sites are interspersed with Pax5 binding sites within the hs5–8 region (5).

## FUNCTIONAL ANALYSIS OF 3' RR REGULATORY ELEMENTS THROUGH TARGETED DELETIONS

### 3' RR ENHANCERS

Analysis of the function of the endogenous 3' RR began with the description of spontaneous 3' RR deletion mutants identified in cell lines. For example, a low-producing variant (LP1.2) of a mouse plasmacytoma cell line was shown to have sustained a deletion of the entire 3' RR (15, 48). This suggested that the 3' RR supported high levels of *Igh* expression in plasma cells. With the development of both transgenic (49, 50) and endogenous models, the 3' RR has been over many years the focus of multiple targeted deletions [reviewed in Ref. (24)]. Although the efficiency of targeting of this 3' RR region has been hampered, perhaps because of its complex structure, there has been gradual, ongoing success. Deletion of individual enhancers had no significant phenotypic consequence implying that the remaining elements, each constellation of which is different, can provide 3' RR function. Deletion of two or more enhancers gave phenotypic consequences of varying degrees, e.g., deletion of hs3b and hs4 together eliminated class switching to all isotypes except for IgG1 (51). Now, there are mice from which the entire ~28 kb 3' enhancer region has been deleted (2), and these have provided a clear demonstration of the potency of the complete 3' RR enhancer region. Without 3' RR enhancers, mice are able to express only reduced levels of IgM at the plasma cell stage, they lack class switch recombination to all isotypes (2) and they are deficient in SHM (3). There is no impairment of VDJ joining (52). Studies by the Cogne laboratory showed that 3' RR enhancers hs1.2 and hs4 were transcriptionally active in B cells, and hs1.2 could be targeted by AID, revealed by detectable although relatively low levels of SHM (30). These AID-dependent mutational and recombination processes involving the 3' RR with S $\mu$  resulted in deletion of the entire *IgC<sub>H</sub>* region and B cell death (30). This revealed an ongoing competition between generation of live class switched mutated B cells and dead B cells, termed “locus

suicide” by the authors. In all, these data strongly show that the 3' RR enhancer region (hs3a–hs4) is critical for CSR and SHM and functions through synergy among the multiple 3' RR enhancers.

### CTCF/Pax5 BINDING REGION

Similar to the analysis of the 3' RR enhancers, we used targeted deletion to examine the effect of the CTCF/Pax5 binding region of the 3' RR on *Igh* expression (53). We were surprised to find that deletion of the 8 kb hs5–7 region resulted only in a mild phenotype. There was an increase in recombination of the most proximal *D* gene, *DQ52*, to *J<sub>H</sub>3*, a reduction in contraction between distal *V<sub>H</sub>J558* and proximal *V<sub>H</sub>J183* genes, and an ~2-fold increase in *V<sub>H</sub>J183* gene usage—all suggesting a modest contribution of the CTCF/Pax5 region of the 3' RR to steps in VDJ joining. Nonetheless, upon targeted deletion of hs5–7, there were essentially normal levels of *Igh* recombination for *V<sub>H</sub>* formation and CSR, normal levels of *Igh* expression and allelic exclusion, and B cell development was unaffected. One possibility to account for these observations was our finding that two CTCF sites remained downstream of the seven sites that had been deleted, in the segment called “38” in the manuscript and now termed hs8. In addition, CTCF sites are associated with each of the downstream non-*Igh* genes. This suggests that a full deletion of CTCF sites in this region might reveal a more extensive phenotype.

## PHYSICAL INTERACTION OF THE 3' RR WITH ITS TARGET SITES IN THE *Igh* LOCUS

*V<sub>H</sub>* promoters and *I* promoters that drive germline transcription (GT) for CSR are situated quite far in a linear distance from the 3' RR; yet it is implied that they all function together through physical interaction (Figure 1). In fact, our finding of an inversion of the *Igh* locus in a variant of the MPC11 plasma cell line that extended from the *V<sub>H</sub>* through to the 3' RR (54) was indicative of a loop formed by interactions between DNA sequences at *V<sub>H</sub>* and 3' RR inversion breakpoints (55). Chromosome conformation capture (3C) technology has been important in documenting interactions that occur in a cellular context, by fixing these by formaldehyde cross-linking, cutting away intervening DNA stretches with restriction enzymes, ligating remaining neighboring fragments, and documenting these interactions by PCR with selected primer pairs. Using 3C, we sought to confirm the dependence of H chain expression in plasma cells on an intact 3' RR (55): indeed, we found that the 3' RR interacted with the *J<sub>H</sub>*–E $\mu$  region. This interaction took place even in cells in which E $\mu$  was deleted. Not only was there interaction between the 3' RR and its target sequences, but there was also interaction among component 3' RR enhancers and insulators, including the CTCF/Pax5 binding unit (hs5–8) (3' CBE). Notably, substitution of hs1.2 by the *NeoR* gene in a variant of the MPC11 plasma cell line resulted in loss of *Igh* expression (56) and abrogation of the 3' RR loop structure; i.e., looping was essential for *Igh* expression. Collectively, these experiments show that the entire 3' RR, including enhancers and insulators, works as a physical unit.

The Kenter laboratory focused on normal spleen cells stimulated to undergo switching for their 3C experiments (57). They reported that in resting B cells, but not in T cells, the 3' RR interacted with the VDJ–E $\mu$  region. Upon LPS  $\pm$  IL4 stimulation of

splenic B cells, they found that specific *I/switch* regions that drive GT were brought into the *VDJ*–3' RR loop. Splenic B cells from mice that were unable to carry out GT and CSR as a result of the combined deletion of the *hs3b* and *hs4* 3' RR enhancers failed to show these interactions.

Interestingly, mice bearing the combined deletion of *hs3a* and *hs3b* (58) had no defects in GT or CSR, but interactions between the 3' RR and *I/switch* regions that ordinarily were cytokine-dependent were already at an induced level in the *hs3a/hs3b* deleted mouse. Collectively, these data provided support for a loop interaction model by which H chain expression and CSR are dependent on physical interaction of the 3' RR with target *Igh* sequences.

Presuming that the CTCF/Pax5 region (3' CBE) of the 3' RR interacts with other *Igh*-associated CTCF sites, such candidate CTCF target sites have been defined by colleagues using array analysis and genome-wide ChIP (59, 60). Moving 3'–5' upstream of the 3' RR, there are no CTCF sites in *C<sub>H</sub>* and *J<sub>H</sub>* genes until those detected in the 5' *D<sub>H</sub>* segment (1, 61). The *V<sub>H</sub>* region contains multiple CTCF sites, some associated with specific Pax5 binding sites, termed PAIR (62). Recent studies showed that the 3' RR CTCF/Pax5 binding region interacts with the two *D<sub>H</sub>*-associated CTCF sites (IGCR1), targeted deletion of which showed their critical role in appropriate regulation of VDJ joining (1). We might, therefore, predict that the deletion of the complete 3' RR CTCF-binding region with which IGCR1 interacts would have a major influence on VDJ joining.

Genome-wide analyses have been used to identify interactions between 3' RR elements, e.g., *hs3b* and *hs8*, and the rest of the *Igh* locus (47, 63). Studies with 4C (47) have identified Pax5-dependent interactions in Rag<sup>−/−</sup> pro-B cells where *V<sub>H</sub>* genes are poised to contract prior to VDJ joining. These 3' RR interactions are maintained even when individual regulatory elements, such as Eμ, IGCR1, and the entire 3' RR enhancer region from *hs3A* to *hs4*, are deleted. This implies some independent means of interaction, perhaps involving retained 3' RR CTCF-binding sites, or synergy among regulatory elements that enables continued interactions even when single elements are deleted. Notably, chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) of long-range chromatin interactions has revealed interactions of the 3' RR with transcribed *Igh* genes in B cells activated by LPS + IL4 that are not detected in embryonic stem (ES) cells (63), in accord with developmental differences in 3' RR interactions.

### WHAT PROTEINS SUPPORT LOOP FORMATION INVOLVING THE 3' RR?

To tackle this question, we used lentiviral delivery of shRNA directed against expression of CTCF, Oct-2, and a coactivator of Oct-2, namely Pou2af1 [i.e., OCAB, OBF-1, BOB1, each of which, as proteins that bound to the 3' RR, was a candidate for loop promotion (64)]. Despite reduced levels of these proteins in response to shRNA, no alterations in loop formation or *Igh* expression were observed in the mouse MPC11 plasmacytoma cell line we examined. Interestingly, in contrast to our observations, there was a report that reduction of *Pou2af1* expression in the same plasma cell line using step-wise selection of a cell line containing two independent shRNA's led to reduction in *Igh* expression and

3C interactions (65). Accordingly, we concluded (64) that there are likely some conditions under which Pou2af1 can facilitate 3C interactions involving the 3' RR, among them the possibility that this approach had selected a variant cell line that was dependent on Pou2af1. Potentially, under other conditions, 3C interactions depend on a protein other than Pou2af1, or on synergy involving more than one protein, or on the low levels of individual proteins remaining from inefficient knock down.

### EPIGENETIC REGULATION OF 3' RR DURING B CELL DEVELOPMENT

Over several years, we have worked to know how components of the 3' RR are individually regulated, enabling them, in turn, to act together as a unit in CSR and SHM, and potentially also for VDJ joining. "Active" DNA segments are generally associated with DNase I hypersensitivity, specific histone marks, and DNA demethylation, which will be discussed individually below.

#### HISTONE MARKS

Non-B cells that were studied had varying profiles of histone acetylation of the 3' RR (66). For example, a macrophage cell line had active marks of AcH3 and AcH4, while T cells lacked both AcH3 and AcH4. In mouse erythroleukemic (MEL) cells, the CTCF-binding region, but not 3' RR enhancers, was associated with acetylated histones. Therefore, active histone marks of the 3' RR were not necessarily B cell-specific. In B cells, modules of the 3' RR sequentially acquire active histone marks during development (44). The CTCF-rich region first acquires these marks, followed progressively 5' to *hs4* and then to the palindromic enhancers. ChIP experiments of the 3' RR showed that in pro-B cells, *hs5* and *hs6* of the CTCF-binding region were associated with AcH4 and low levels of AcH3, while *hs4* was also associated with AcH4 but not with AcH3. In pre-B cells, the entire *hs4*–8 region was associated with both acetylated marks; and then in B cells, the palindromic enhancers also acquired these marks. These observations suggest step-wise activation of different modules of the 3' RR during B cell development, raising the possibility that specific combinations of 3' RR modules, involving palindromic enhancers, *hs4*, and the CTCF-binding region, have specific functional contributions.

#### DNA DEMETHYLATION

Early studies had bluntly monitored DNase I hypersensitivity and DNA demethylation in the region now shown to contain the entire 3' RR (12), but as the complete 3' RR structure became known, a finer analysis was made possible (66). The CTCF-binding region was generally constitutively demethylated in all cell types analyzed. In several sources of non-B cells, the 3' RR's palindromic region was demethylated without a corresponding association with active histone marks. However, in B cells, three epigenetic marks – DNase I hypersensitivity, active histone marks, and DNA demethylation – were collectively engaged; and progressive demethylation paralleled acquisition of active histone marks. *hs4* and downstream CTCF-binding sites were DNase I hypersensitive and demethylated, as assessed by relative sensitivity to the methylation-sensitive isoschizomers *HpaII* and *MspI*, beginning in pro-B cells and extending throughout B cell development. The palindromic region became hypersensitive and partially demethylated only later in B cell development. We found upon comparison

of wild-type with Pax5-deficient pro-B cells in which Pax5 expression could be reinitiated, that in the absence of Pax5, there was scattered demethylation of the palindromic region. Re-expression of Pax5 could promote methylation of the palindromic region. These findings suggested that Pax5 was a critical factor in over-all B cell-specific epigenetic regulation of the 3' RR. In other studies involving targeted deletion of linker histone *H1* genes, we found that linker histone H1 was also important for the methylation of hs4–hs8 in wild-type ES cells.

### EPIGENETIC REGULATION OF 3' RR DURING CSR

Despite the critical role of the 3' RR in CSR (and SHM), there are no apparent changes in histone marks in the 3' RR during switching in cultured cells (44). This implies that the 3' RR in resting B cells is already epigenetically poised for its activity for CSR. Instead, we have observed dynamic changes in Pax5 interaction over time in response to LPS stimulation (5). In resting B cells, Pax5 bound hs4 and the 3' RR was mostly methylated. When GT was at a peak at ~48 h after commencement of LPS stimulation, Pax5 binding had shifted from hs4 upstream to hs1.2 and downstream to hs7. By 72 h, when CSR was essentially complete, Pax5 had resumed its beginning position at hs4. ChIP analysis of cell sources that were deficient in GT or in CSR showed differences in these Pax5 binding patterns, in accord with the notion that shifts in Pax5 binding reflected mechanisms by which the 3' RR supported GT and CSR. We have proposed a model by which mouse 3' RR enhancers form a scaffold through which Pax5 can interact. Deletion of any individual enhancer leaves residual Pax5 sites in each of the remaining enhancers and in the CTCF/Pax5 binding region, which allows the 3' RR to remain functional.

### ON THE HORIZON: EXPERIMENTS ON THE 3' RR

#### HOW DOES THE 3' RR FUNCTION?

The 3' RR enhancer region is critical for GT and CSR, and SHM, and the CTCF/Pax5 binding region could contribute to VDJ joining. We predict that a scaffold formed by modules of the 3' RR supports physical interactions with target sequences required to accomplish these various activities. With deletions of individual 3' RR enhancers having little phenotypic consequence, one can ask how many different structural solutions are there to 3' RR activity? Why are there multiple modules? Do the changes in epigenetic alterations of 3' RR modules that occur during development indicate specific activities for individual modules? Could the 3' RR help target DNA repair proteins involved in SHM or CSR? What roles does the 3' RR share with other cis acting sequences that are critical for SHM, such as those in the light chain loci? What is distinctive about CSR, which is specific to the *Igh* locus? What is the role of the conserved palindrome? How did the 3' RR evolve? What are the species-specific aspects of 3' RR regulation?

Recent experiments on the  $\beta$ -globin LCR have identified hierarchical regulation by multiple transcription factors (67). Binding of individual factors can provide a foundation for subsequent binding of other factors. Experiments of these types on the 3' RR could be equally informative in answering how this region functions. Indeed, more complete deletion of 3' RR CTCF-binding sites, and targeted deletions and mutations in 3' RR modules would also be

informative. The new CRISPR technology (68) should facilitate these constructions and provide answers to many questions.

### IS THE CTCF/Pax5 BINDING REGION THE TERMINUS OF B CELL-SPECIFIC REGULATORS OF THE *Igh* LOCUS?

A persuasive set of experiments says “yes” to the role of the 3' RR as a terminus of *Igh* regulation via chromatin accessible marks. These experiments have shown that active chromatin marks extend unilaterally 3'–5' from the 3' RR as far as 450 kb when *Igh-myc* translocations are assayed in endemic Burkitt lymphoma samples (69). This supports the identification of the CTCF/Pax5 region as a functional insulator of the *Igh* locus. In addition, 4C studies have implicated hs8 as a 3' boundary for *Igh* locus interactions (47). Yet, as described below, there is a replicative terminus further downstream, raising the possibility of additional *Igh* locus regulators.

### ROLE OF REPLICATIVE TERMINUS DOWNSTREAM OF CTCF/Pax5 REGION

Our experiments in collaboration with the Schildkraut laboratory identified an origin of an ~500 kb *Igh* temporal replicative transition region in MEL (non-B) cells. DNA replication initiates ~45 kb downstream of the CTCF/Pax5 module of the 3' RR between *crip1* and *Tmem121* (Figure 1) and extends progressively 3'–5' throughout S phase to replicate the 3' RR, *C<sub>H</sub>*, *J<sub>H</sub>*, *D<sub>H</sub>*, and most proximal *V* regions (70, 71). All *V<sub>H</sub>* genes replicate late in S phase. In pre-B cells, the entire *Igh* locus replicates early in S phase, indicating the firing of multiple origins that are ordinarily quiescent in non-B cells (71). In B cells, a temporal transition region is again apparent, but origins appear to be closer to or within the 3' RR, suggesting that the replication landmark is flexible (72). These data implied that *Igh* replication is under B cell-specific developmental control. In MEL cells, the downstream origin, which is located beyond the limits of the 3' RR, may be a terminus for *Igh* locus regulation. It is of interest that changes in *Igh* DNA replication are associated with changes in nuclear location of the *Igh* locus (71, 73) but can be independently regulated (74). In pro- and pre-B cells, the *Igh* locus is located away from the nuclear periphery, while in MEL and ES cells, and in B and plasma cells, the *Igh* locus is located at the nuclear periphery. These observations raise the question of whether there are finer demarcations of nuclear subcompartments generally associated with the *Igh* locus and the 3' RR? What regulates the movement of the locus from one position to another?

### IS THE 3' RR INVOLVED IN INTER-CHROMOSOMAL INTERACTIONS?

The mechanism by which recurrent translocations involving the *Igh* locus take place and the role of the 3' RR are under close scrutiny (75, 76). Epner and colleagues have reported a role for the 3' RR in transvection involving allelic interactions (77). Further, our studies have identified a region between hs4 and hs5 that has a methylation signature indicative of allelic expression (66). The Skok laboratory has observed allelic interactions in *Igh* genes, which are evident during steps of VDJ joining (78). The various 3C technologies and their broader counterparts, as noted in part above (47, 63), should be very informative about the contribution of the 3' RR to genetic domains of interaction.

## IS THERE A ROLE OF THE 3' RR AS A SUPER ENHANCER?

Recent genome-wide studies have reported “super enhancers” (79, 80), DNA segments substantially larger than other “enhancer” regions and identified as having strong binding sites for BRD4, a member of the bromodomain and extraterminal (BET) sub-family of human bromodomain proteins, and for the Mediator complex with which BRD4 interacts. By these criteria, the 3' RR was predicted to be a super enhancer in multiple myeloma cells (79), where it upregulates expression of the *myc* oncogene to which it is juxtaposed as a result of a chromosomal translocation. An inhibitor of BRD4, JQ1, can lead to downregulation of *myc* expression in multiple myeloma cells. However, *myc* appears to be suppressed by JQ1 regardless of whether it is associated with *Igh* sequences through translocation (81), potentially via B cell-specific enhancers of *myc* (47, 63). Is the 3' RR a super enhancer? Under what circumstances? Does the 3' RR share features in common with other “super enhancers”? How might the 3' RR become a super enhancer?

## ACKNOWLEDGMENTS

This review focuses primarily on work carried out in my laboratory with the support of NIH RO1AI13509 and RO1 AI41572, and summarizes studies by many individuals. I thank Sandra Giannini, Jennifer Michaelson, Mallika Singh, Fang Liao, Nancy Martinez, Nasrin Ashouian, Charles-Felix Calvo, Alexa Price-Whelan, Chaoqun Chen, Francine Garrett-Bakelman, Alejandro Sepulveda, Rabih Hassan, Vincenzo Giambra, Steven Gordon, Alexander Emelyanov, Zhongliang Ju, Sanjukta Chatterjee, and Sabrina Volpi, together with other former students and post-doctoral fellows, who provided important foundation for these studies. I thank our collaborators at Einstein, Carl Schildkraut, Randall Little, Olga Ermakova, Jie Zhou, Qiaoxin Yang, Winfried Edelmann, Harry Hou, Uwe Werling, Britta Will, Ulrich Steidl, Matthew Scharff, and Sergio Roa: our colleagues elsewhere, Laurel Eckhardt, Clifford Snapper, Victor Lobanenko, Dmitry Loukinov, Roy Riblet, Fumi Matsuda, Domenico Frezza, Ann Feeney, and Jiyoti Verma-Gaur; and other colleagues who have studied the 3' regulatory region, especially Michel Cogne, John Manis, Amy Kenter, Fred Alt, Wesley Dunnick, and Rafael Casellas. I especially thank Xiaohua Wang, Boris Bartholdy, and Shanzhi Wang for their critical comments and discussion on this manuscript. I regret not being able to cite all those who have contributed to the fascinating studies of the *Igh* locus.

## REFERENCES

- Guo C, Yoon HS, Franklin A, Jain S, Ebert A, Cheng HL, et al. CTCF-binding elements mediate control of V(D)J recombination. *Nature* (2011) **477**:424–30. doi:10.1038/nature10495
- Vincent-Fabert C, Fiancette R, Pinaud E, Truffinet V, Cogne N, Cogne M, et al. Genomic deletion of the whole *Igh* 3' regulatory region (hs3a, hs1.2, hs3b, hs4) dramatically affects class switch recombination and Ig secretion to all isotypes. *Blood* (2010) **116**:1895–8. doi:10.1182/blood-2010-01-264689
- 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**:1501–7. doi:10.1084/jem.20130072
- Michaelson JS, Singh M, Snapper CM, Sha WC, Baltimore D, Birshtein BK. Regulation of 3' *Igh* enhancers by a common set of factors, including kappa B-binding proteins. *J Immunol* (1996) **156**:2828–39.
- Chatterjee S, Ju Z, Hassan R, Volpi SA, Emelyanov AV, Birshtein BK. Dynamic changes in binding of immunoglobulin heavy chain 3' regulatory region to protein factors during class switching. *J Biol Chem* (2011) **286**:29303–12. doi:10.1074/jbc.M111.243543
- Max EE, editor. *Immunoglobulins: Molecular Genetics*. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins (2008).
- Eckhardt LA, Birshtein BK. Independent immunoglobulin class-switch events occurring in a single myeloma cell line. *Mol Cell Biol* (1985) **5**:856–68.
- Zaller DM, Eckhardt LA. Deletion of a B-cell-specific enhancer affects transfected, but not endogenous, immunoglobulin heavy-chain gene expression. *Proc Natl Acad Sci U S A* (1985) **82**:5088–92. doi:10.1073/pnas.82.15.5088
- Pettersson S, Cook GP, Bruggemann M, Williams GT, Neuberger MS. A second B cell-specific enhancer 3' of the immunoglobulin heavy-chain locus. *Nature* (1990) **344**:165–8. doi:10.1038/344165a0
- Vincent-Fabert C, Fiancette R, Cogne M, Pinaud E, Denizot Y. The *Igh* 3' regulatory region and its implication in lymphomagenesis. *Eur J Immunol* (2010) **40**:3306–11. doi:10.1002/eji.201040778
- Liebersohn R, Giannini SL, Birshtein BK, Eckhardt LA. An enhancer at the 3' end of the mouse immunoglobulin heavy chain locus. *Nucleic Acids Res* (1991) **19**:933–7. doi:10.1093/nar/19.4.933
- Giannini SL, Singh M, Calvo C-F, Ding G, Birshtein BK. DNA regions flanking the mouse *Ig* 3' alpha enhancer are differentially methylated and DNase I hypersensitive during B cell differentiation. *J Immunol* (1993) **150**:1772–80.
- Matthias P, Baltimore D. The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the alpha constant region. *Mol Cell Biol* (1993) **13**:1547–53.
- Madisen L, Groudine M. Identification of a locus control region in the immunoglobulin heavy-chain locus that deregulates c-myc expression in plasmacytoma and Burkitt's lymphoma cells. *Genes Dev* (1994) **8**:2212–26. doi:10.1101/gad.8.18.2212
- Michaelson JS, Giannini SL, Birshtein BK. Identification of 3' alpha-hs4, a novel *Ig* heavy chain enhancer element regulated at multiple stages of B cell differentiation. *Nucleic Acids Res* (1995) **23**:975–81. doi:10.1093/nar/23.6.975
- Zhou J, Ashouian N, Delepine M, Matsuda F, Chevillard C, Riblet R, et al. The origin of a developmentally regulated *Igh* replicon is located near the border of regulatory domains for *Igh* replication and expression. *Proc Natl Acad Sci U S A* (2002) **99**:13693–8. doi:10.1073/pnas.212392399
- Chen C, Birshtein BK. A region of the 20 bp repeats lies 3' of human *Ig* Calpha1 and Calpha2 genes. *Int Immunol* (1996) **8**:115–22. doi:10.1093/intimm/8.1.115
- Chen C, Birshtein BK. Virtually identical enhancers containing a segment of homology to murine 3' *Igh*-E(hs1.2) lie downstream of human *Ig* C alpha 1 and C alpha 2 genes. *J Immunol* (1997) **159**:1310–8.
- Pinaud E, Aupetit C, Chauveau C, Cogne 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**:2981–5. doi:10.1002/eji.1830271134
- Mills FC, Harindranath N, Mitchell M, Max EE. Enhancer complexes located downstream of both human immunoglobulin Calpha genes. *J Exp Med* (1997) **186**:845–58. doi:10.1084/jem.186.6.845
- Birshtein BK, Chen C, Saleque S, Michaelson JS, Singh M, Little RD. Murine and human 3' *Igh* regulatory sequences. *Curr Top Microbiol Immunol* (1997) **224**:73–80.
- Khamlichi AA, Pinaud E, Decourt C, Chauveau C, Cogne M. The 3' *Igh* regulatory region: a complex structure in a search for a function. *Adv Immunol* (2000) **75**:317–45. doi:10.1016/S0065-2776(00)75008-5
- Cogne M, Birshtein BK. Regulation of class switch recombination. In: Honjo T, Alt FW, Neuberger M, editors. *Molecular Biology of B Cells*. San Diego, CA: Elsevier Academic Press (2004). p. 289–305.
- Pinaud E, Marquet M, Fiancette R, Peron 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
- Chauveau C, Cogne M. Palindromic structure of the *Igh* 3' locus control region. *Nat Genet* (1996) **14**:15–6. doi:10.1038/ng0996-15
- Saleque S, Singh M, Little RD, Giannini SL, Michaelson JS, Birshtein BK. Dyad symmetry within the mouse 3' *Igh* regulatory region includes two virtually identical enhancers (C alpha3'E and hs3). *J Immunol* (1997) **158**:4780–7.
- Sepulveda MA, Garrett FE, Price-Whelan A, Birshtein BK. Comparative analysis of human and mouse 3' *Igh* regulatory regions identifies distinctive



- structural features. *Mol Immunol* (2004) **42**:605–15. doi:10.1016/j.molimm.2004.09.006
28. D'Addabbo P, Scascitelli M, Giambra V, Rocchi M, Frezza D. Position and sequence conservation in amniota of polymorphic enhancer HS1.2 within the palindrome of *Igh* 3' regulatory region. *BMC Evol Biol* (2011) **11**:71. doi:10.1186/1471-2148-11-71
  29. Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, et al. PipMaker – a web server for aligning two genomic DNA sequences. *Genome Res* (2000) **10**:577–86. doi:10.1101/gr.10.4.577
  30. Peron 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**:931–4. doi:10.1126/science.1218692
  31. Saleque S, Singh M, Birshtein BK. Ig heavy chain expression and class switching in vitro from an allele lacking the 3' enhancers DNase I-hypersensitive hs3A and hs1.2. *J Immunol* (1999) **162**:2791–803.
  32. Giambra V, Martinez-Labarga C, Giufre M, Modiano D, Simpori J, Gisladotir BK, et al. Immunoglobulin enhancer HS1.2 polymorphism: a new powerful anthropogenetic marker. *Ann Hum Genet* (2006) **70**:946–50. doi:10.1111/j.1469-1809.2006.00273.x
  33. Frezza D, Tolusso B, Giambra V, Gremese E, Marchini M, Nowik M, et al. Polymorphisms of the *Igh* enhancer HS1.2 and risk of systemic lupus erythematosus. *Ann Rheum Dis* (2012) **71**:1309–15. doi:10.1136/ard.2010.147025
  34. Kim A, Dean A. Chromatin loop formation in the beta-globin locus and its role in globin gene transcription. *Mol Cells* (2012) **34**:1–5. doi:10.1007/s10059-012-0048-8
  35. Liao F, Giannini SL, Birshtein BK. A nuclear DNA-binding protein expressed during early stages of B cell differentiation interacts with diverse segments within and 3' of the *Ig H* chain gene cluster. *J Immunol* (1992) **148**:2909–17.
  36. Singh M, Birshtein BK. NF-HB (BSAP) is a repressor of the murine immunoglobulin heavy-chain 3' alpha enhancer at early stages of B-cell differentiation. *Mol Cell Biol* (1993) **13**:3611–22.
  37. Barberis A, Widenhorn K, Vitelli L, Busslinger M. A novel B-cell lineage-specific transcription factor present at early but not late stages of differentiation. *Genes Dev* (1990) **4**:849–59. doi:10.1101/gad.4.5.849
  38. Singh M, Birshtein BK. Concerted repression of an immunoglobulin heavy-chain enhancer, 3' alpha E(hs1.2). *Proc Natl Acad Sci U S A* (1996) **93**:4392–7. doi:10.1073/pnas.93.9.4392
  39. Sepulveda MA, Emelyanov AV, Birshtein BK. NF-kappaB and Oct-2 synergize to activate the human 3' *Igh* hs4 enhancer in B cells. *J Immunol* (2004) **172**:1054–64.
  40. Duan H, Heckman CA, Boxer LM. The immunoglobulin heavy-chain gene 3' enhancers deregulate bcl-2 promoter usage in t(14;18) lymphoma cells. *Oncogene* (2007) **26**:2635–41. doi:10.1038/sj.onc.1210061
  41. Duan H, Xiang H, Ma L, Boxer LM. Functional long-range interactions of the *Igh* 3' enhancers with the bcl-2 promoter region in t(14;18) lymphoma cells. *Oncogene* (2008) **27**:6720–8. doi:10.1038/nc.2008.286
  42. Kim EC, Edmonston CR, Wu X, Schaffer A, Casali P. The HoxC4 homeodomain protein mediates activation of the immunoglobulin heavy chain 3' hs1.2 enhancer in human B cells. Relevance to class switch DNA recombination. *J Biol Chem* (2004) **279**:42258–69. doi:10.1074/jbc.M407496200
  43. Birshtein BK. The role of CTCF binding sites in the 3' immunoglobulin heavy chain regulatory region. *Front Genet* (2012) **3**:251. doi:10.3389/fgene.2012.00251
  44. Garrett FE, Emelyanov AV, Sepulveda MA, Flanagan P, Volpi S, Li F, et al. 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* (2005) **25**:1511–25. doi:10.1128/MCB.25.4.1511-1525.2005
  45. Nicolas L, Chaudhuri J. 4C-ing the *Igh* landscape. *Immunity* (2013) **39**:199–201. doi:10.1016/j.immuni.2013.08.014
  46. Ohlsson R, Renkawitz R, Lobanenko V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet* (2001) **17**:520–7. doi:10.1016/S0168-9525(01)02366-6
  47. Medvedovic J, Ebert A, Tagoh H, Tamir IM, Schwickert TA, Novatchkova M, et al. Flexible long-range loops in the VH gene region of the *Igh* locus facilitate the generation of a diverse antibody repertoire. *Immunity* (2013) **39**:229–44. doi:10.1016/j.immuni.2013.08.011
  48. Gregor PD, Morrison SL. 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* (1986) **6**:1903–16.
  49. Dunnick WA, Collins JT, Shi J, Westfield G, Fontaine C, Hakimpour P, et al. Switch recombination and somatic hypermutation are controlled by the heavy chain 3' enhancer region. *J Exp Med* (2009) **206**:2613–23. doi:10.1084/jem.20091280
  50. Dunnick WA, Shi J, Zerbato JM, Fontaine CA, Collins JT. Enhancement of antibody class-switch recombination by the cumulative activity of four separate elements. *J Immunol* (2011) **187**:4733–43. doi:10.4049/jimmunol.1101808
  51. 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**:187–99. doi:10.1016/S1074-7613(01)00181-9
  52. Rouaud P, Vincent-Fabert C, Fiancette R, Cogne M, Pinaud E, Denizot Y. Enhancers located in heavy chain regulatory region (hs3a, hs1.2, hs3b, and hs4) are dispensable for diversity of VDJ recombination. *J Biol Chem* (2012) **287**:8356–60. doi:10.1074/jbc.M112.341024
  53. Volpi SA, Verma-Gaur J, Hassan R, Ju Z, Roa S, Chatterjee S, et al. Germline deletion of *Igh* 3' regulatory region elements hs 5, 6, 7 (hs5-7) affects B cell-specific regulation, rearrangement, and insulation of the *Igh* locus. *J Immunol* (2012) **188**:2556–66. doi:10.4049/jimmunol.1102763
  54. Calvo C-F, Giannini SL, Martinez N, Birshtein BK. DNA sequences 3' of the *Igh* chain cluster rearrange in mouse B cell lines. *J Immunol* (1991) **146**:1353–60.
  55. Ju Z, Volpi SA, Hassan R, Martinez N, Giannini SL, Gold T, et al. Evidence for physical interaction between the immunoglobulin heavy chain variable region and the 3' regulatory region. *J Biol Chem* (2007) **282**:35169–78. doi:10.1074/jbc.M705719200
  56. Lieberman R, Ong J, Shi X, Eckhardt LA. Immunoglobulin gene transcription ceases upon deletion of a distant enhancer. *EMBO J* (1995) **14**:6229–38.
  57. Wuerffel R, Wang L, Grigera F, Manis J, Selsing E, Perlot T, et al. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity* (2007) **22**:711–22. doi:10.1016/j.immuni.2007.09.007
  58. Yan Y, Pieretti J, Ju Z, Wei S, Christin JR, Bah F, et al. Homologous elements hs3a and hs3b in the 3' regulatory region of the murine immunoglobulin heavy chain (*Igh*) locus are both dispensable for class-switch recombination. *J Biol Chem* (2011) **286**:27123–31. doi:10.1074/jbc.M111.230995
  59. Degner SC, Wong TP, Jankevicius G, Feeney AJ. Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J Immunol* (2009) **182**:44–8.
  60. Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, et al. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the *Igh* locus and antisense transcription in pro-B cells. *Proc Natl Acad Sci U S A* (2011) **108**:9566–71. doi:10.1073/pnas.1019391108
  61. Featherstone K, Wood AL, Bowen AJ, Corcoran AE. The mouse immunoglobulin heavy chain V-D intergenic sequence contains insulators that may regulate ordered V(D)J recombination. *J Biol Chem* (2010) **285**:9327–38. doi:10.1074/jbc.M109.098251
  62. Ebert A, McManus S, Tagoh H, Medvedovic J, Salvaggio G, Novatchkova M, et al. The distal v(h) gene cluster of the *Igh* locus contains distinct regulatory elements with pax5 transcription factor-dependent activity in pro-B cells. *Immunity* (2011) **34**:175–87. doi:10.1016/j.immuni.2011.02.005
  63. Kieffer-Kwon KR, Tang Z, Mathe E, Qian J, Sung MH, Li G, et al. Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell* (2013) **155**:1507–20. doi:10.1016/j.cell.2013.11.039
  64. Ju Z, Chatterjee S, Birshtein BK. Interaction between the immunoglobulin heavy chain 3' regulatory region and the *Igh* transcription unit during B cell differentiation. *Mol Immunol* (2011) **49**:297–303. doi:10.1016/j.molimm.2011.08.024
  65. Ren X, Siegel R, Kim U, Roeder RG. Direct interactions of OCA-B and TFII-I regulate immunoglobulin heavy-chain gene transcription by facilitating enhancer-promoter communication. *Mol Cell* (2011) **42**:342–55. doi:10.1016/j.molcel.2011.04.011
  66. Giambra V, Volpi S, Emelyanov AV, Pflugh D, Bothwell AL, Norio P, et al. 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* (2008) **28**:6123–33. doi:10.1128/MCB.00233-08



67. Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* (2012) **149**:1233–44. doi:10.1016/j.cell.2012.03.051
68. Menke DB. Engineering subtle targeted mutations into the mouse genome. *Genesis* (2013) **51**:605–18. doi:10.1002/dvg.22422
69. Kovalchuk AL, Ansarah-Sobrinho C, Hakim O, Resch W, Tolarova H, Dubois W, et al. Mouse model of endemic Burkitt translocations reveals the long-range boundaries of Ig-mediated oncogene deregulation. *Proc Natl Acad Sci U S A* (2012) **109**:10972–7. doi:10.1073/pnas.1200106109
70. Ermakova OV, Nguyen LH, Little RD, Chevillard C, Riblet R, Ashouian N, et al. Evidence that a single replication fork proceeds from early to late replicating domains in the *Igh* locus in a non-B cell line. *Mol Cell* (1999) **3**:321–30. doi:10.1016/S1097-2765(00)80459-1
71. Zhou J, Ermakova OV, Riblet R, Birshtein BK, Schildkraut CL. Replication and subnuclear location dynamics of the immunoglobulin heavy-chain locus in B-lineage cells. *Mol Cell Biol* (2002) **22**:4876–89. doi:10.1128/MCB.22.13.4876-4889.2002
72. Demczuk A, Gauthier MG, Veras I, Kosiyatrakul S, Schildkraut CL, Buslinger M, et al. Regulation of DNA replication within the immunoglobulin heavy-chain locus during B cell commitment. *PLoS Biol* (2012) **10**:e1001360. doi:10.1371/journal.pbio.1001360
73. Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, et al. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* (2002) **296**:158–62. doi:10.1126/science.1068768
74. Zhou J, Saleque S, Ermakova O, Sepulveda MA, Yang Q, Eckhardt LA, et al. Changes in replication, nuclear location, and expression of the *Igh* locus after fusion of a pre-B cell line with a T cell line. *J Immunol* (2005) **175**:2317–20.
75. Gostissa M, Yan CT, Bianco JM, Cogne M, Pinaud E, Alt FW. Long-range oncogenic activation of *Igh*-c-myc translocations by the *Igh* 3' regulatory region. *Nature* (2009) **462**:803–7. doi:10.1038/nature08633
76. Zhang Y, McCord RP, Ho YJ, Lajoie BR, Hildebrand DG, Simon AC, et al. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* (2012) **148**:908–21. doi:10.1016/j.cell.2012.02.002
77. Liu H, Huang J, Wang J, Jiang S, Bailey AS, Goldman DC, et al. Transvection mediated by the translocated cyclin D1 locus in mantle cell lymphoma. *J Exp Med* (2008) **205**:1843–58. doi:10.1084/jem.20072102
78. Hewitt SL, Yin B, Ji Y, Chaumeil J, Marszalek K, Tenthorey J, et al. RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. *Nat Immunol* (2009) **10**:655–64. doi:10.1038/ni.1735
79. Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* (2013) **153**:320–34. doi:10.1016/j.cell.2013.03.036
80. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* (2013) **153**:307–19. doi:10.1016/j.cell.2013.03.035
81. Chapuy B, McKeown MR, Lin CY, Monti S, Roemer MG, Qi J, et al. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* (2013) **24**:777–90. doi:10.1016/j.ccr.2013.11.003

**Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 January 2014; accepted: 27 March 2014; published online: 21 April 2014.

Citation: Birshtein BK (2014) Epigenetic regulation of individual modules of the immunoglobulin heavy chain locus 3' regulatory region. *Front. Immunol.* 5:163. doi: 10.3389/fimmu.2014.00163

This article was submitted to *B Cell Biology*, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Birshtein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Oct2 and Obf1 as facilitators of B:T cell collaboration during a humoral immune response

Lynn Corcoran<sup>1,2\*</sup>, Dianne Emslie<sup>1,2</sup>, Tobias Kratina<sup>1,2</sup>, Wei Shi<sup>1,2</sup>, Susanne Hirsch<sup>1,2</sup>, Nadine Taubenheim<sup>1,2</sup> and Stephane Chevrier<sup>1,2†</sup>

<sup>1</sup> Molecular Immunology Division, Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia

<sup>2</sup> Department of Medical Biology, The University of Melbourne, Melbourne, VIC, Australia

## Edited by:

Ananda L. Roy, Tufts University School of Medicine, USA

## Reviewed by:

Mikael Sigvardsson, Linköping University, Sweden

Bruce David Mazer, Montreal Children's Hospital, Canada

## \*Correspondence:

Lynn Corcoran, Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Melbourne, VIC 3052, Australia

e-mail: corcoran@wehi.edu.au

## †Present address:

Stephane Chevrier, Institute of Molecular Life Sciences, University of Zurich, Zurich, Switzerland

The Oct2 protein, encoded by the *Pou2f2* gene, was originally predicted to act as a DNA binding transcriptional activator of immunoglobulin (Ig) in B lineage cells. This prediction flowed from the earlier observation that an 8-bp sequence, the “octamer motif,” was a highly conserved component of most Ig gene promoters and enhancers, and evidence from over-expression and reporter assays confirmed Oct2-mediated, octamer-dependent gene expression. Complexity was added to the story when Oct1, an independently encoded protein, ubiquitously expressed from the *Pou2f1* gene, was characterized and found to bind to the octamer motif with almost identical specificity, and later, when the co-activator Obf1 (OCA-B, Bob.1), encoded by the *Pou2af1* gene, was cloned. Obf1 joins Oct2 (and Oct1) on the DNA of a subset of octamer motifs to enhance their transactivation strength. While these proteins variously carried the mantle of determinants of Ig gene expression in B cells for many years, such a role has not been borne out for them by characterization of mice lacking functional copies of the genes, either as single or as compound mutants. Instead, we and others have shown that Oct2 and Obf1 are required for B cells to mature fully *in vivo*, for B cells to respond to the T cell cytokines IL5 and IL4, and for B cells to produce IL6 normally during a T cell dependent immune response. We show here that Oct2 affects *Syk* gene expression, thus influencing B cell receptor signaling, and that Oct2 loss blocks *Slamf1* expression *in vivo* as a result of incomplete B cell maturation. Upon IL4 signaling, Stat6 up-regulates Obf1, indirectly via Xbp1, to enable plasma cell differentiation. Thus, Oct2 and Obf1 enable B cells to respond normally to antigen receptor signals, to express surface receptors that mediate physical interaction with T cells, or to produce and respond to cytokines that are critical drivers of B cell and T cell differentiation during a humoral immune response.

**Keywords: Oct2, Obf1, Syk, Slamf1, B:T collaboration, cytokines**

## INTRODUCTION

Octamer binding protein 2, or Oct2, is encoded by the *Pou2f2* gene. It was one of the first cell type-specific transcription factors identified and cloned (1). As indicated by its name, it is a founding member of a family of DNA binding proteins concurrently discovered, that share a conserved bipartite DNA binding domain comprising a homeobox-like domain and a second conserved sequence entitled the POU domain, for the *Pit1*, *Oct1/Oct2*, *Unc86* proteins (2). Oct2 binds to a conserved consensus DNA sequence, the “octamer motif” found in the promoters and enhancers of many genes, including those encoding immunoglobulins (3, 4). The Obf1 protein encoded by the *Pou2af1* gene, which is also known as OCA-B and Bob.1 was subsequently cloned using a yeast 1-hybrid screen for B cell proteins that physically interact with Oct1 or Oct2 (5–7). While Oct1/Oct2 and Obf1 share the capacity to bind to and activate genes adjacent to octamer motifs, they are selective in the genes to which they bind. The selectivity of target gene binding is determined, in part, by the sequence of the octamer motif, and whether it conforms to one of two classes of

site, designated “PORE” and “MORE” motifs (8). Whether binding mediates activation or repression is also influenced by the participation of cofactors [reviewed by Tantin (9)], including Obf1, which can potentiate the transactivation potential of Oct1 and Oct2 (8, 10).

Oct2 is expressed primarily but not exclusively in the B cell lineage, where it increases with cellular activation (11). Neurons, macrophages, and T cells have also been shown to express *Oct2* (12–18). Oct2 is required for post-natal survival (19), so must regulate critically important genes outside of the immune system. These will not be discussed here. The *Oct2* gene is large, displays complex splicing patterns, and encodes protein isoforms with multiple essential activation domains (20–22). Oct2 is largely localized to the nucleus. *Obf1* expression is mostly restricted to B lineage cells, where it is also highly induced upon activation (23). Zwilling et al. (24) have reported expression in T cells, but myeloid cells do not express *Obf1* (15). A small protein of ~35 kDa, Obf1 is found in both the nucleus and cytoplasm, where a proportion may be tethered to the cell membrane after post-translational myristoylation

(25), and a potential role for membrane-associated Obf1 in B cell receptor (BCR) signaling has been proposed (26).

A series of studies have shown that Oct2 and Obf1 are required for full functional and phenotypic maturation of B cells. In single knockout (KO) mice of each gene, peripheral B cells are numerically reduced and display some features of immature transitional cells (27, 28). The peritoneal B1 and splenic marginal zone (MZ) populations are missing in *Oct2*<sup>-/-</sup> mice (27, 29). *Obf1*<sup>-/-</sup> mice are viable and fertile, but show B cell developmental defects (30, 31), have an expanded B1 cell population (32). They also lack MZ B cells (33) and completely fail to produce germinal centers (GCs), the sites of cognate B cell:T cell interaction and expansion, upon immunization, or infection (34–37). Both Oct2- and Obf1-deficient splenic B cells display aberrant responses to BCR signaling and other characteristics of immature B cells (27, 34, 38). Oct2-deficient B cells also fail to respond to lipopolysaccharide (LPS), which signals through TLR4 (38). *In vivo*, serum immunoglobulin (Ig) levels in both mutants, particularly those that are T cell dependent, are strongly reduced (34–36, 38). Mice doubly deficient for Oct2 and Obf1 show a stronger humoral deficiency phenotype, reflecting the distinct activities of the two factors, but still express Ig genes (39). Thus, the two factors are not required, singly or in combination, for Ig gene expression by B cells.

Detailed functional studies on Oct2- and Obf1-deficient B cells *in vitro* and *in vivo* have identified a number of genes regulated by the two factors. Oct2 directly regulates the gene encoding CD36, a class B scavenger receptor family (40), but only in B cells, not in macrophages or dendritic cells (41–43). However, no role for CD36 in B cells has been determined (44). Oct2-deficient B cells have been shown to be defective in their responses to the T cell cytokine IL5 as a result of the direct regulation of the *Cd125* gene encoding the IL5R $\alpha$  chain (29). IL5 promotes antibody-secreting cell (ASC) differentiation in mouse B cells (45), and *Oct2*<sup>-/-</sup> B cells are defective in this process (29). In another study, it was shown that both Oct2 and Obf1 contribute to the regulation of IL6 production by activated B cells, through direct effects, at least by Oct2, on the *Il6* gene (11). As Obf1 does not contact DNA (5), it is difficult using current procedures to prove direct interaction of Obf1 with putative target gene loci. IL6 is important during T follicular helper (Tfh) cell polarization (46). We have also shown, using the same quantitative tools that identified the role of Oct2 and IL5 in ASC differentiation, that Obf1 is required for T cell dependent ASC differentiation, but not isotype switching, both *in vitro* and *in vivo* (29).

In addition to these established roles for Oct2 and Obf1 in B cells, we include below data from studies on other genes that we have found to be differentially regulated in Oct2- and Obf1-deficient B cells. These include the genes encoding the Syk protein, which is an important transducer of BCR signals and *Slamf1*, an essential mediator of cell:cell contact, especially in the context of a developing GC. Expression of Syk and *Slamf1* are sensitive to Oct2 loss, through different mechanisms. We also show that Obf1 is downstream of Stat6 in the IL4 signaling pathway of B cells, with Xbp1, another Stat6 target, its direct activator. We include these data to add to our understanding of the valuable roles that Oct2 and Obf1 play in B cell responses to antigen and to T cell help.

## RESULTS

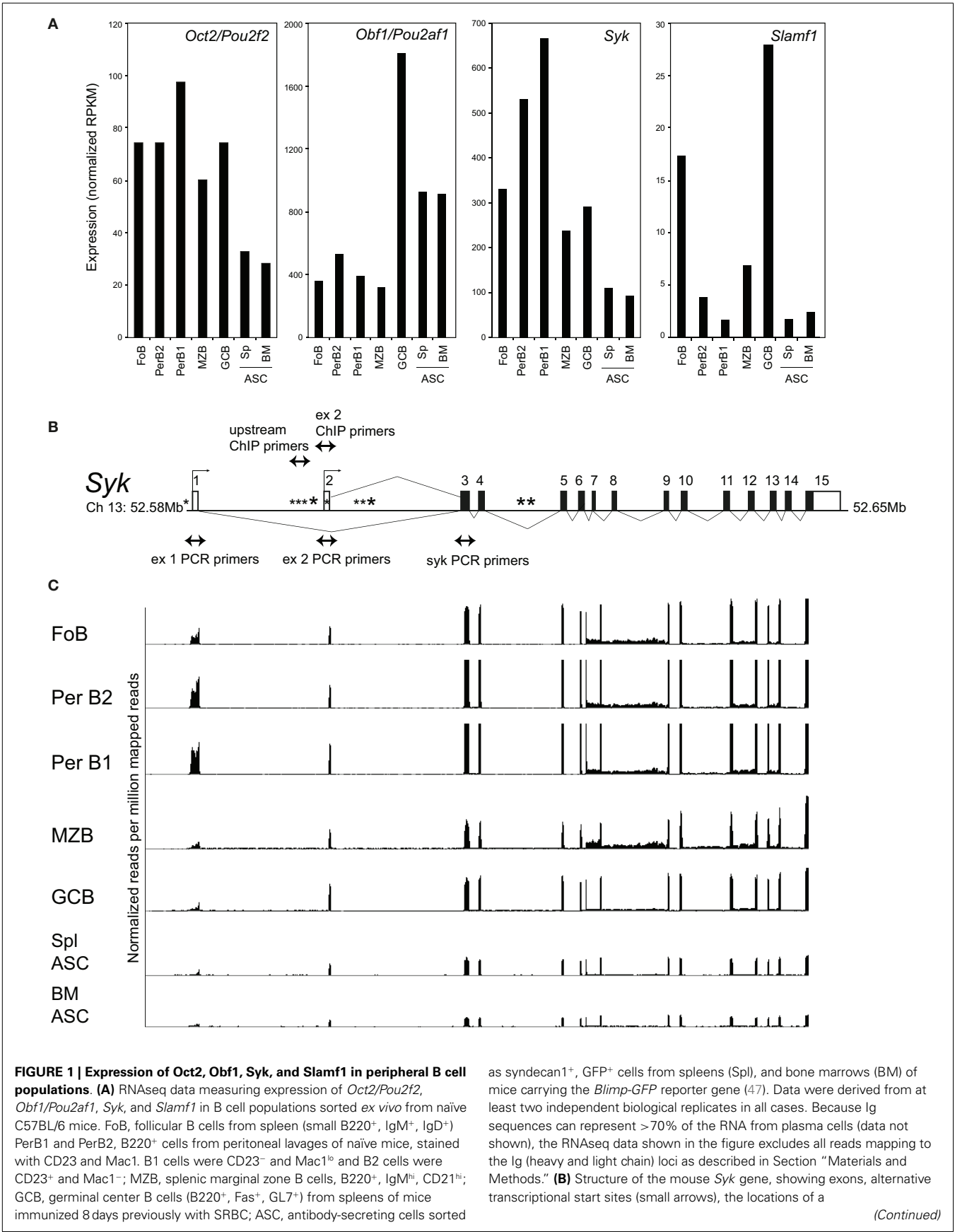
Consistent with their distinct roles *in vivo*, Oct2 and Obf1 have quite distinct patterns of expression in peripheral B cells, as measured by RNAseq of sorted populations from naive C57BL/6 mice (**Figure 1A**). Oct2 levels are highest in B1 cells of the peritoneal cavity, and decline with terminal differentiation to ASCs. In contrast, Obf1 levels peak in GC B cells, which require Obf1 for their generation, and remain high in ASC. For contrast, expression of *Syk* and *Slamf1* in the same populations are shown in **Figure 1A**, as these two genes are influenced directly and indirectly, respectively, by Oct2, and will be discussed below.

### Oct2 MODULATES B CELL RECEPTOR SIGNALING BY FINE-TUNING Syk EXPRESSION

A microarray screen for Oct2-dependent genes identified the tyrosine kinase *Syk* as a potential target gene. Characterization of the murine *Syk* promoter using 5' RACE identified two alternative transcriptional initiation sites (**Figure 1B**) that append alternative 5' non-coding exons to *Syk* mRNAs in B cells. Both transcripts encode the same protein, as the start of Syk translation lies in an exon common to the two transcripts. RNAseq data for *Syk* in sorted B cell populations show that the usage of exons 1 or 2 varies subtly among them (**Figure 1C**).

The promoter upstream of exon 2 is positively regulated by Oct2. B cells sorted from *Oct2*<sup>-/-</sup> mice (**Figure 2A**) have a lower level of *Syk* transcripts and protein than wild type (WT) mice (**Figures 2B,C**). Using qPCR to distinguish *Syk* transcripts derived from exon 1 or exon 2, we found that those derived from exon 2 were selectively reduced in *Oct2*<sup>-/-</sup> B cells (**Figure 2D**). To confirm the influence of Oct2 on the *Syk* exon 2 promoter, we stably introduced an estradiol-inducible form of Oct2 (29) into a cloned *Oct2*<sup>-/-</sup> B lymphoma cell line, OM1 (48). Upon treatment with estradiol, there was no effect on expression from exon 1 with either the vector only control or the inducible Oct2 construct. However, estradiol induction of Oct2 selectively enhanced *Syk* expression from exon 2 and culminated in markedly increased *Syk* mRNA levels (**Figures 2E,F**). A DNA sequence search revealed three perfect consensus octamer sequences in the *Syk* gene (**Figure 1B**). Chromatin immunoprecipitation (ChIP) of Oct2 on DNA from the WT B lymphoma WEHI231 showed strong enrichment of DNA from the promoter of the known Oct2 target gene *Cd36* (**Figure 2G**). The DNA adjacent to the perfect octamer sequence upstream of exon 2 was also significantly enriched in the ChIP, but adjacent sequences in intron 1/2 were not. Thus, Oct2 can directly increase *Syk* levels in B cells, acting at one of two alternative promoters.

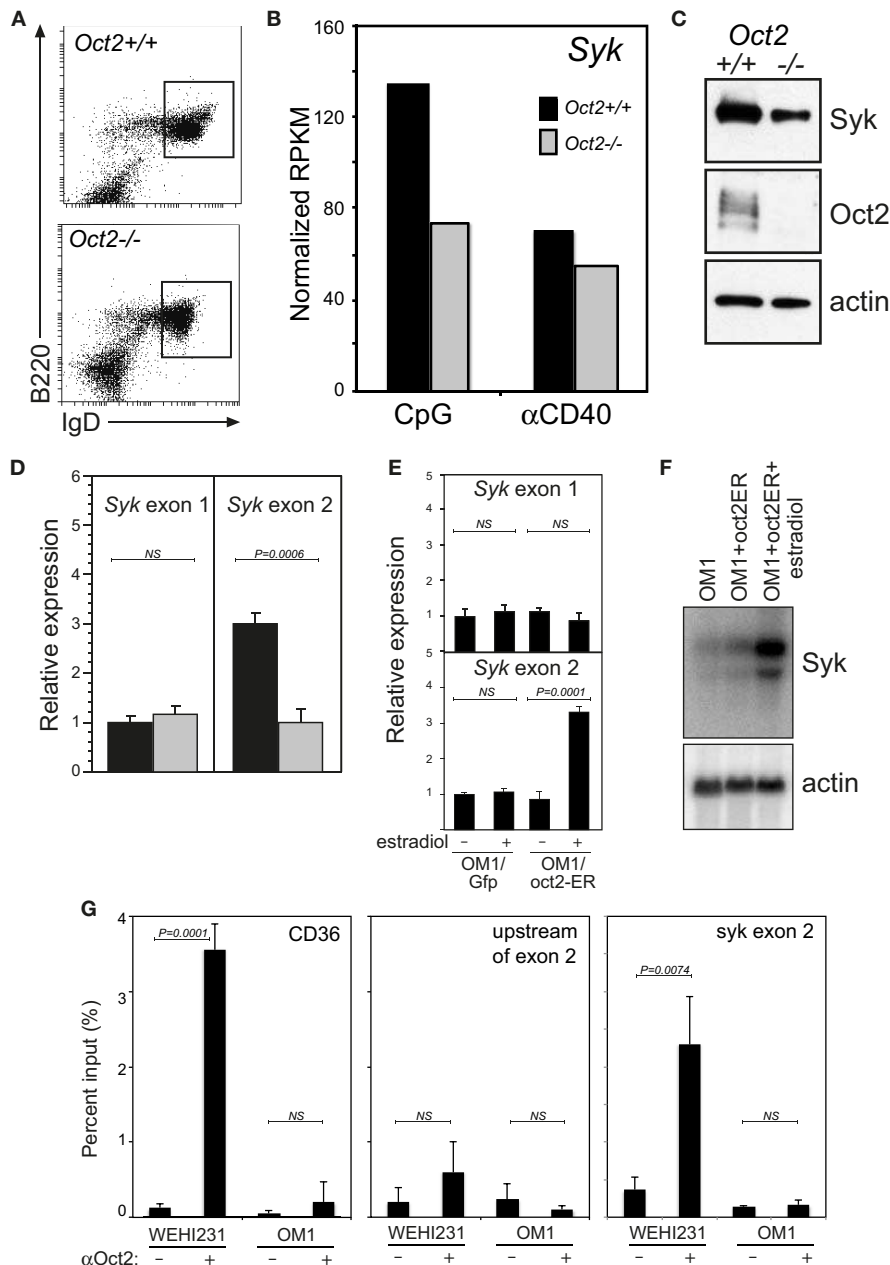
As mentioned above, Oct2 is required for the full functional and phenotypic maturation of B cells, such that peripheral B cells in *Oct2*<sup>-/-</sup> mice are numerically reduced and display some features of immature transitional cells, and the B1 and MZ populations are missing (27, 29, 38). Like transitional B cells, Oct2-deficient B cells are killed rather than activated by BCR cross-linking [Ref. (38, 49); see also **Figure 3A**]. However, *Oct2*<sup>-/-</sup> B cells respond normally to the survival factor Baff/BlyS, and survive to expand normally when BCR signaling occurs in the presence of Baff (**Figures 3A,B** and data not shown). *Oct2*<sup>-/-</sup> mice expressing a *Bcl2* transgene still lack B1 and MZ B cells (data not shown). Thus abnormal



**FIGURE 1 | Continued**

perfect consensus octamer motif (\*) and the positions of PCR primers used here. Filled boxes indicate protein coding sequence, and open boxes, sequence comprising the 5' and 3' untranslated regions of *Syk* mRNA. **(C)** RNAseq tracks showing expression of the *Syk* gene exons in different sorted

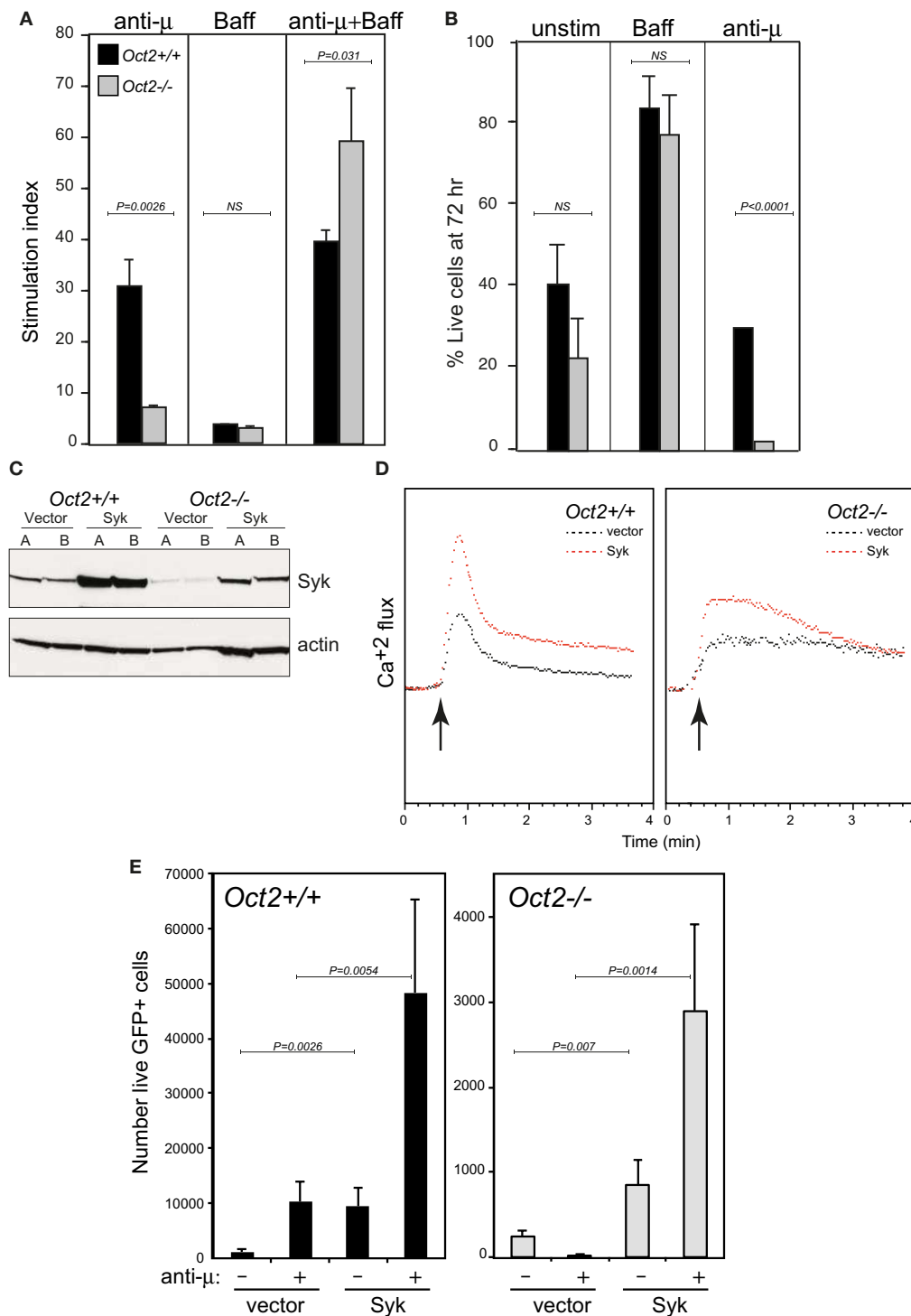
B cell populations, normalized to library size, and aligned with the gene structure of **(B)**. Note that exon 1, as shown in this panel, is not included in the RefSeq (Mouse mm9, July 2007) map of *Syk* mRNA, but is represented in alternate *Syk* transcripts ENSMUST00000120135 and ENSMUST00000118756 in the Ensembl database.



**FIGURE 2 | Oct2 directly and selectively activates transcription from *Syk* exon 2.** **(A)** Splenic B cells from WT and *Oct2* KO mice stained for B220 and IgD expression and sorted for phenotypically mature Fo B cells. **(B)** *Syk* RNAseq data from WT (black bars) and *Oct2*<sup>-/-</sup> (gray bars) B cells, sorted as in **(A)**, activated for 48 h with CpG or anti-CD40. **(C)** *Syk* protein in sorted resting Fo B cells from *Oct2*<sup>+/+</sup> or *Oct2*<sup>-/-</sup> mice. **(D)** qPCR of *Syk* mRNA distinguishing transcripts initiated at exons 1 or 2 in sorted splenic Fo B from WT and *Oct2* KO mice. Expression is relative to that of the *hmbs* housekeeping gene. Values are means  $\pm$  SD of triplicate assays. **(E)** Specific induction of transcription from *Syk* exon 2 upon Oct2 over-expression in

OM1 cells, which are *Oct2*<sup>-/-</sup> (48), as shown by qPCR. Values are means  $\pm$  SD of triplicates. **(F)** Northern blot for total *Syk* mRNA from a parallel experiment. **(G)** Quantitation of Oct2 chromatin immunoprecipitation (ChIP) qPCR data, showing enrichment of the *Cd36* promoter [a known Oct2 target gene (41, 42)] and *Syk* sequences upstream of exon 2. Oct2:DNA complexes were precipitated from WEHI231 B lymphoma cells (which are *Oct2*<sup>+/+</sup>) and OM1 B lymphoma cells. Oct2 does not bind appreciably to adjacent sequences in *Syk* intron 1/2. Values are means  $\pm$  SD of triplicate assays. *P* values were calculated using the unpaired Student's *t*-test; NS, not significant.





**FIGURE 3 | Oct2 is required for normal signaling from the B cell receptor, and ectopic Syk expression enhances the response in both WT and KO B cells. (A)** B cell proliferation in response to BCR signaling, Baff, or both in combination after 3 days. Stimulation index is calculated as proliferation relative to unstimulated cells. Filled bars, Oct2<sup>+/+</sup>; gray bars, Oct2<sup>-/-</sup>. All values are the mean of triplicates  $\pm$  SD. **(B)** Survival, assessed by propidium iodide exclusion, of B cells cultured with Baff or with anti- $\mu$  for 3 days. Filled bars, Oct2<sup>+/+</sup>; gray bars, Oct2<sup>-/-</sup>. All values are the mean of triplicates  $\pm$  SD. **(C)** Syk protein levels in cloned B lymphoma cells

transduced with a Syk-expressing retrovirus. The Oct2<sup>+/+</sup> and Oct2<sup>-/-</sup> cell lines are BC1 and OM1, respectively (48). **(D)** Cytometric measurement of Ca<sup>2+</sup> flux in clones of BC1 and OM1 cells transduced with vector only (black) or a Syk-expressing retrovirus (red). The arrow indicates the timing of addition of anti- $\mu$  to cross-link the BCR. **(E)** Primary splenic B cells from Oct2<sup>+/+</sup> and Oct2<sup>-/-</sup> mice were activated, transduced (see Materials and Methods) and subsequently treated with anti- $\mu$ . The number of live, transduced (GFP<sup>+</sup>) cells in each culture after 48 h is shown. Values are means ( $n=4$ )  $\pm$  SD. Filled bars, Oct2<sup>+/+</sup>; gray bars, Oct2<sup>-/-</sup>.

survival properties are not responsible for the lack of these two populations in the Oct2 mutant mice.

We speculated that the reduced Syk levels in *Oct2*<sup>-/-</sup> B cells might contribute to their failure to mature *in vivo* and respond to BCR signals *in vitro*. We constructed a retroviral vector expressing Syk, and infected and cloned WT (BC1) and Oct2-deficient (OM1) lymphoma cells (Figure 3C). Boosting Syk protein levels enhanced the BCR response, as measured by calcium flux, in both WT and *Oct2*<sup>-/-</sup> cells (Figure 3D). Finally, using transduction of primary B cells (see Materials and Methods), we found that elevating Syk levels improved the proliferation of WT B cells, both unstimulated, and more strongly, upon BCR cross-linking, and that *Oct2*<sup>-/-</sup> B cells complemented with Syk retrovirus were activated to expand, rather than be killed by a BCR signal (Figure 3E). The rescue was not complete, as cell survival was still lower overall in the mutant cell cultures. This is likely to reflect technical limitations of the assay, including comparative infectivity of WT and mutant cells, and the correct timing of exogenous Syk expression in the context of the BCR signal. However, the results strongly suggest that Syk levels are limiting in *Oct2*<sup>-/-</sup> and, to a lesser extent, in WT B cells. We propose that Oct2 regulates Syk gene expression to enable positive selection through the BCR and therefore entrance to the mature follicular B cell pool, and it may similarly enable differentiation of B1 and MZ B cells, which are highly dependent on BCR signal strength.

#### Oct2 INDIRECTLY AND SELECTIVELY REGULATES *Slamf1* EXPRESSION ON B CELLS

*Slamf1* encodes CD150, a lymphocyte signaling and adhesion molecule (50) that is expressed in B cells and T cells and at very low levels in myeloid cells (see www.immgen.org). Our RNAseq analysis of activated B cells indicated that *Slamf1* was expressed at abnormally low levels in Oct2 KO B compared to controls (Figure 4A). This was confirmed by flow cytometric analysis (Figure 4B). Resting WT and LPS-activated B cells expressed similar levels of *Slamf1*, but levels on Oct2-deficient B cells were much lower than controls under both conditions. Activated WT T cells up-regulated *Slamf1* from resting levels. However, *Slamf1* was not Oct2-dependent in resting or activated T cells, or in macrophages expanded from fetal liver, despite Oct2 normally being expressed in these cell types (15, 16, 18).

We next asked whether *Slamf1* was a direct Oct2 target using cloned WT (WEHI231) and KO (OM1) B lymphoma lines transduced with the inducible Oct2-ER vector, as described above. As expected, the Oct2 target *Cd36* gene was expressed in WEHI231 but not OM1 (Figure 4C). Estradiol treatment enhanced *Cd36* levels in WEHI231, and strongly induced *Cd36* expression in OM1 cells. However, *Slamf1* expression, while low in the *Oct2*<sup>-/-</sup> line, was not increased by Oct2 induction. Therefore, Oct2 is unlikely to directly regulate *Slamf1* transcription in B cells.

Examination of the pattern of expression of the *Slamf1* gene during B cell development, using the Immgen database (<http://www.immgen.org>) and by flow cytometry of bone marrow (BM) and peripheral B cell populations indicated that *Slamf1* is a marker of B cell maturation, appearing during the transition from immature (IgM<sup>hi</sup>/IgD<sup>lo</sup>) to mature (IgM<sup>lo</sup>/IgD<sup>hi</sup>) follicular B cells of the spleen (Figure 4D), with MZ B cells expressing intermediate

*Slamf1* levels. We conclude that loss of Oct2 blocks B cell maturation before the *Slamf1*<sup>+</sup> stage, thereby indirectly regulating its expression.

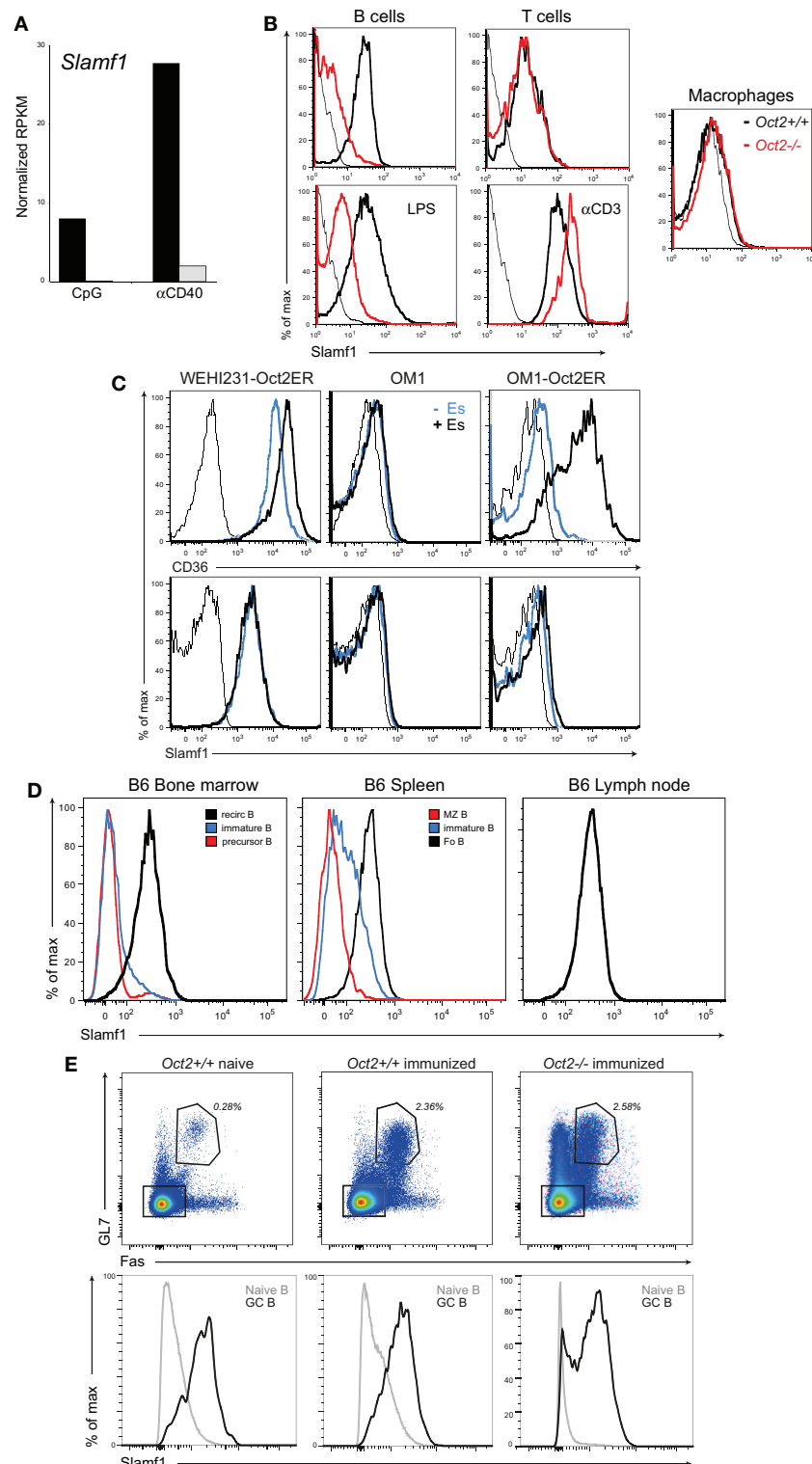
It has been shown that B cells lacking *Slamf1* cannot form the lasting interactions with Tfh cells that are required for GC formation (51), and yet Oct2 mice do form GC upon infection and immunization (11). We immunized WT and KO mice with SRBC and stained GC cells for *Slamf1* 9 days later. *Oct2*<sup>+/+</sup> and *Oct2*<sup>-/-</sup> GC B cells expressed similar levels of *Slamf1*. This indicates that Oct2 is dispensable for *Slamf1* expression, and that endogenous signals driving the GC response override the Oct2 maturation defect *in vivo* (Figure 4E). To explore the nature of these signals, we tested the capacity of a number of mitogens and cytokines to induce *Slamf1* expression on Oct2-null B cells *in vitro*. Anti-μ, anti-CD40, Baff, and IL4, tested singly or in all possible combinations, failed to induce appreciable *Slamf1* on *Oct2*<sup>-/-</sup> B cells, suggesting that other factors and cells contribute *in vivo*. Collectively these studies point to an Oct2-regulated differentiation step that enables efficient FoB cell maturation.

#### Obf1 ENABLES T CELL DEPENDENT ASC DIFFERENTIATION DRIVEN BY IL4

We have shown that both Oct2 and Obf1 affect a B cell's capacity to differentiate to ASC in response to particular cytokines, with Oct2 regulating the response to IL5 (29), and Obf1 being essential for ASC differentiation driven by IL4 (52).

We determined that Obf1 lies downstream of Stat6 in the IL4 signaling cascade (Figures 5A–C). Interestingly, while the IL4/Stat6 axis drives both isotype switching and ASC differentiation (45), Obf1 is dispensable for IL4 driven switching (52). In order to learn how IL4 and Stat6 regulate *Obf1* expression, we performed microarray analysis on the WEHI231 B lymphoma, an IL4 responsive line. Because Stat6 exists in a latent form in the cytoplasm, and is activated by phosphorylation to enter the nucleus and act as a transcription factor (53), direct Stat6 targets would be activated after IL4 signaling even in the absence of new protein synthesis. Cells therefore were treated with cycloheximide (CHX) to inhibit translation, and 1 h later, IL4 was added. A parallel culture did not receive IL4. After another 4 h, RNA was prepared from both. Analysis using Illumina Sentrix Mouse v1.1 arrays identified a small number of genes whose expression rose two to fourfold during this short period of IL4 stimulation, genes likely to be regulated directly by Stat6. Among these were three transcription factors: Nfil3, Vdr, encoding the 1,25-dihydroxyvitamin D3 receptor, and Xbp1. Their induction under these conditions was validated by qPCR (Figure 5D). Interestingly, *Obf1* transcription was not elevated by IL4 treatment in the presence of CHX, indicating that it is not a direct Stat6 target gene (Figure 5D).

As *Obf1* RNA is elevated by IL4 treatment in B cells in the absence of CHX (Figure 5B), we hypothesized that one of the three transcription factors directly regulated by Stat6 might drive *Obf1* expression. We therefore constructed ER fusion vectors for each to determine their effects on Obf1 expression. The Xbp1-ER expression vector contained the mature, processed, and active form of this factor (54). Clones of each line were cultured for 24 h unstimulated, with estradiol to induce each fusion protein, or



**FIGURE 4 | Oct2 influences and Slamf1 expression in B cells. (A)** Analysis of Slamf1 expression in Oct2-null and Oct2<sup>+/+</sup> B cells activated *in vitro*. Cells were cultured for 48 h in the presence of CpG or αCD40 before RNA was prepared for RNA sequencing. Filled bars, Oct2<sup>+/+</sup>; gray bars, Oct2<sup>-/-</sup>. **(B)** Slamf1/CD150 protein expression in cells of the indicated genotypes. Cells were assessed directly *ex vivo* (resting, top panels), or were activated

*in vitro* for 48 h with either LPS (B cells) or anti-CD3 (T cells). Macrophages were expanded from fetal liver as described in Section “Materials and Methods.” In the histograms, WT cells are represented with heavy black lines and Oct2<sup>-/-</sup> as red. Unstained controls are indicated by thin black lines.

**(C)** WEHI231 cells (Oct2<sup>+/+</sup>) and OM1 cells (Oct2<sup>-/-</sup>), either uninfected or (Continued)

**FIGURE 4 | Continued**

transduced with an Oct-ER expression vector, were cultured in the presence or absence of estradiol (Es) for 24 h and CD36 and Slamf1 levels were determined by flow cytometry. Thin lines indicate background fluorescence of unstained controls, blue lines represent CD36 and Slamf1 levels in uninoculated cultures, and heavy black lines, the levels after Es induction. **(D)** Slamf1 expression during B cell maturation. Colors indicate the populations represented in each histogram. For bone marrow, recirculating B cells were B220<sup>+</sup> and IgM<sup>+</sup>, immature B were B220<sup>+</sup> and

IgM<sup>+</sup>, and precursor B were B220<sup>+</sup> and IgM<sup>+</sup>. For spleen, immature B cells were IgM<sup>hi</sup>, IgD<sup>lo</sup>, MZ B cells were IgM<sup>hi</sup>, CD21<sup>hi</sup>, and Fo B cells were IgM<sup>lo</sup>, IgD<sup>hi</sup>, CD23<sup>+</sup>, and CD21<sup>+</sup>. Lymph node B cells were IgM<sup>lo</sup>, IgD<sup>hi</sup>, and CD23<sup>+</sup>. **(E)** Flow cytometric analysis of splenic B cells from naïve and immunized mice (9 days after SRBC immunization). Top panels show the percentage of GL7<sup>+</sup> Fas<sup>+</sup> GC B cells (gated) among total B220<sup>+</sup> cells in spleens from mice reconstituted with WT or *Oct2*<sup>-/-</sup> fetal liver. Bottom panels show the Slamf1 levels on non-GC (gray lines) and GC (black line) B cells for each animal, gated as shown in the upper panels.

with IL4. IL4 caused an increase in *Obf1* RNA levels in all cases, as expected. However, estradiol induction only increased *Obf1* levels significantly in the Xbp-ER expressing cell line (**Figure 5E**). Finally, primary splenic B cells from *Xbp1*<sup>+/+</sup>, heterozygous and conditional KO mice were stimulated for 48 h with CD40 ligand and IL4, and protein extracts prepared. Western blots showed that Xbp1-null B cells express markedly less Obf1 than the controls (**Figure 5F**). These data indicate that Nfil3 and Vdr do not influence *Obf1* expression, but that Xbp1 has the capacity to directly activate the *Obf1* gene in response to IL4/Stat6 signaling in B cells.

## DISCUSSION

The data we present here adds to a growing view of Oct2 and Obf1 as essential contributors to the sensing capacity of B cells (**Figure 6**). These two factors enhance the cell's ability to deliver a BCR signal to drive maturation, or to sense a foreign antigen and become activated. We show here that Oct2 may do so by fine-tuning Syk levels. Since Oct2 loss blocks peripheral B cell maturation, such that MZ B cells are missing and mature FoB reduced, Oct2 may play its most important role prior to a divergence point of Fo and MZ B cells (55). The immature Slamf1<sup>+</sup>, Baffr<sup>+</sup>, CD23<sup>+</sup>, CD21<sup>+</sup>, IgM<sup>hi</sup>, D<sup>lo</sup>, HSA<sup>hi</sup> phenotype of *Oct2*<sup>-/-</sup> B cells does not neatly fit into the phenotypic transition of immature to mature B cells (55), and may represent a normally transient phase of B cell maturation. Obf1 is required for normal B cell maturation and MZ B cell development (33), but as Obf1 does not influence *Syk* expression, it is likely that the two factors act through unique subsets of target genes for this aspect of B cell development. Indeed, it has been reported that Obf1 positively regulates *SpiB* gene expression (56), and that *SpiB* is required for normal B cell maturation and BCR signaling, through regulation of *c-rel* (57, 58).

T cell dependent antibody responses depend upon Oct2 and Obf1 in several ways. Both Oct2 and Obf1 are required for B cells to produce normal levels of IL6, a cytokine important in Tfh maturation in the context of a Slamf1-mediated B cell:Tfh cell interaction (11, 59, 60). These molecular interactions enable the initiation of a fruitful T cell-dependent humoral immune response. For example, IL6 produced by activated B cells early in the GC response reinforces the early dendritic cell signal that initiates Tfh differentiation, and IL21 produced by the nascent Tfh enhances both Tfh function and B cell differentiation into GC B cells and ASC (61–66). Slamf1, which is indirectly dependent upon Oct2, is required to prolong the B:Tfh interaction while these important signals are exchanged. Subsequently, Oct2 and Obf1 enable B cells to respond to other Th cell cytokines that drive ASC differentiation. Oct2 regulates *Cd125* expression (29), and so expression of the high

affinity receptor for IL5, an ASC differentiation factor (45). *Obf1* is required for a normal B cell response to IL4, a growth, survival, isotype switching and differentiation factor for B and plasma cells [(45, 67) and this study].

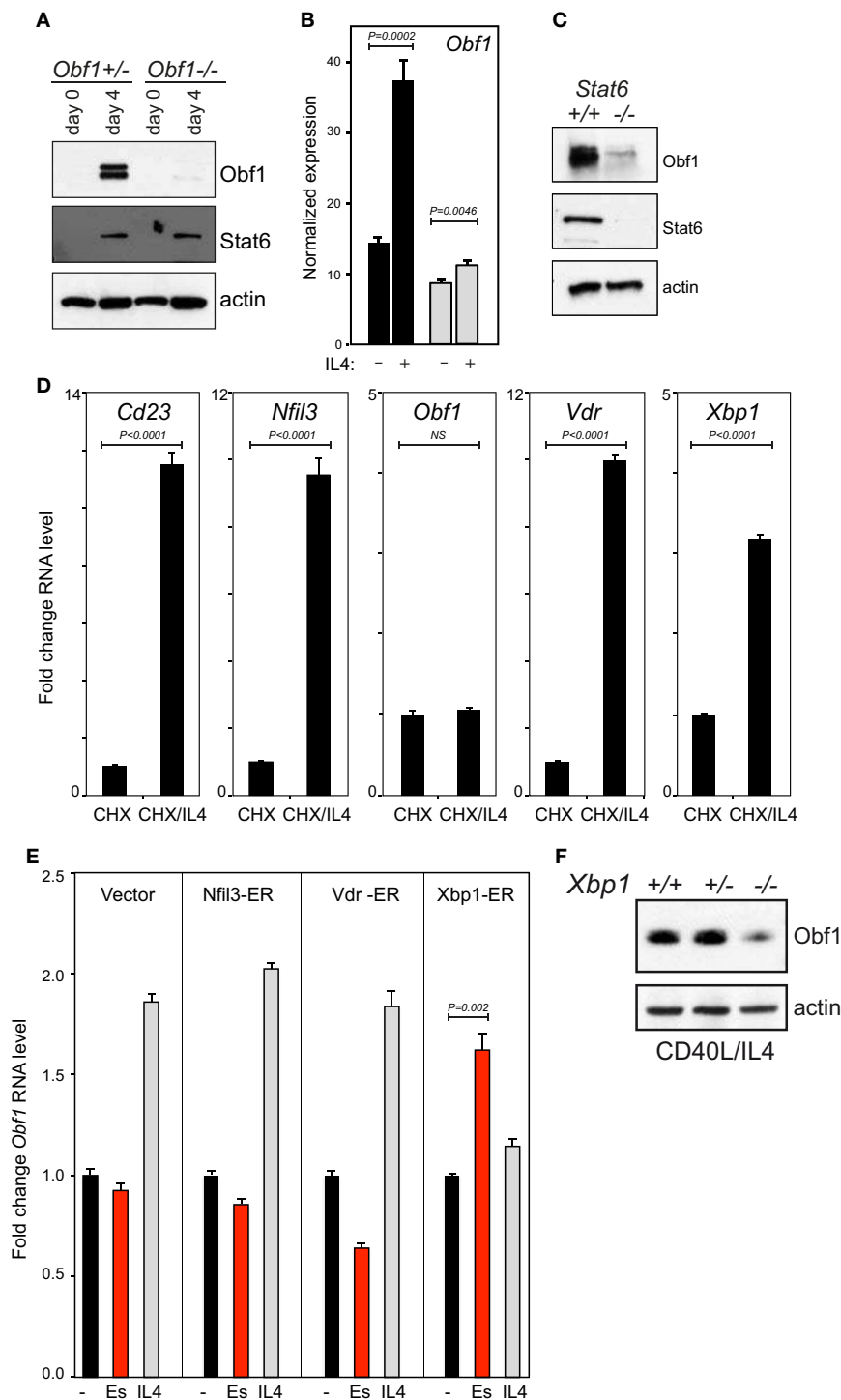
Except for the specific case of Obf1 and a subset of V<sub>L</sub> genes (68), there is no evidence that Oct2 or Obf1 are required for Ig gene expression or for antibody secretion in ASC generated from single or double KO mice (29, 39, 52). Interestingly, however, we show here that Xbp1, which is highly expressed in ASC and known to be required for ASC function [(69–72) and our unpublished data], can directly activate *Obf1*. Shen and Hendershot (73) have also shown in ASC that *Obf1* is a direct target of Xbp1. Accordingly, *Obf1* expression is elevated in normal ASC compared to B cells, unlike Oct2 expression, which declines with differentiation (**Figure 1**). Xbp1, normally associated with the unfolded protein response in ASC, is an IL4 response gene in B cells. We confirm here that *Obf1* is an Xbp1 target gene in FoB cells, and that IL4 and Stat6 directly induce Xbp1. Iwakoshi et al. (70) have earlier shown that, in the context of ASC differentiation, IL4 strongly induces *Xbp1* expression and Stat6 is required. Thus, the selective ASC differentiation defect in Obf1 null B cells under T cell-dependent conditions may reflect an important Stat6-Xbp1-Obf1 axis.

Instead of direct influences on Ig gene expression, poor humoral immune responses in Oct2 or Obf1 mutant mice more likely reflects a paucity of differentiated peripheral B cell populations and weak to absent influences of T cells on these B cells. *Oct2* and *Obf1* are most highly expressed in GC B cells, with Obf1 being essential for their differentiation, but Oct2 dispensable (11). In both mutants, the B cells defects are cell intrinsic. Outstanding questions remain: what are the genes that Oct2 and Obf1 regulate to ensure full peripheral maturation and competent BCR signaling? What are the Obf1 regulated genes that are essential for GC development under all circumstances tested so far, including immunization, infection, and autoimmunity (11, 32, 37)? While these critical Obf1 target(s) are not yet known, candidates such as Bcl6 (or its co-repressor MTA3), Bach2, Irf4 and CXCR5, all critical for normal GC formation or maintenance [see Ref. (74)], can be excluded, as they are not influenced by Obf1 loss (our unpublished data). Ongoing genome-wide RNA expression analysis, coupled with ChIPseq, will enable detailed characterization of the full Oct2 and Obf1 gene regulatory networks, and their shared and unique responsibilities in delivering effective B cell immunity.

## MATERIALS AND METHODS

### CELL LINES, CELL CULTURE, AND RETROVIRAL TRANSDUCTION

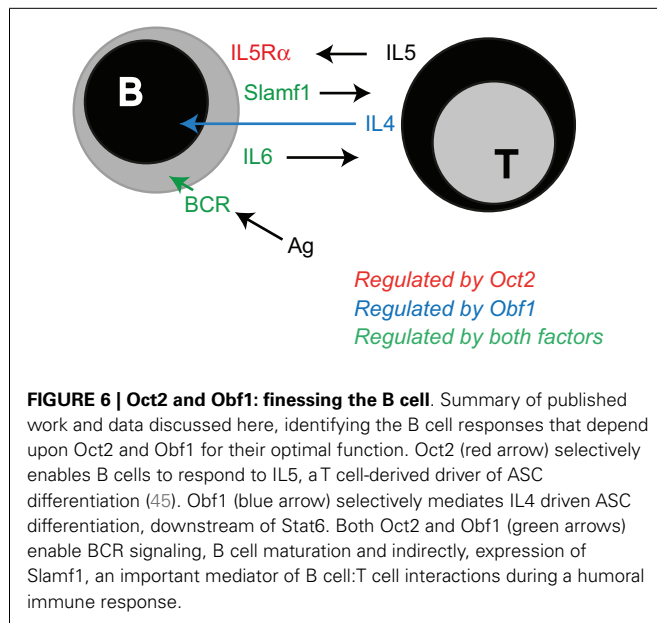
B lymphoma cell lines used here were all generated in house: WEHI231 (75), BC1, and OM1 (48). Primary splenic B and T



**FIGURE 5 | Obf1 is downstream of Stat6 in the IL4 signaling cascade.** (A) Western blots to detect Obf1 and Stat6 proteins in purified B cells of the indicated genotypes, harvested 48 h after anti-CD40 ± IL4 treatment. Actin is included as a loading control. (B) RNA sequencing data showing *Obf1* expression 48 h after anti-CD40 ± IL4 treatment in WT (black bars) and *Stat6*<sup>-/-</sup> B cells (gray bars). (C) Western blot on the same samples as in (B), to detect Stat6 and Obf1 proteins. (D) Induction of IL4 responsive genes in the presence of the translational inhibitor cycloheximide (CHX), as measured by qPCR. (E) Induction of *Obf1* RNA

levels in B lymphoma cell lines each expressing a ER fusion of each of the transcriptional regulators that are direct IL4 target genes: *Nfil3*, *Vdr*, and *Xbp1*. Cells were untreated or treated with either estradiol or IL4 for 24 h. In (B,D,E), values are means ± SD for triplicate samples. (F) Western blot to detect *Obf1* protein in purified B cells from *Xbp1*<sup>+/+</sup>, heterozygous, or conditional KO mice. Cells were cultured in anti-CD40 plus IL4 for 48 h before analysis. In (B,D,E), means ± SD for triplicate measurements are shown, with *P* values determined using the unpaired Student's *t*-test. NS, not significant.





cells were purified using anti-B220 or anti-CD4 microbeads (Miltenyi), as described (11). *Oct2*<sup>+/+</sup> and *Oct2*<sup>-/-</sup> macrophages were expanded *in vitro* from E13 fetal liver using M-CSF, as described (15). Retroviral transduction in all cases used the pMX-pie vector and was performed using spin infection as described (48, 76). For lymphoma lines, cells were cloned post-infection as single GFP<sup>+</sup> cells in puromycin-supplemented medium.

For the retroviral complementation experiment of **Figure 3E**, primary splenic B220<sup>+</sup> cells were stimulated for 24 h with CpG [1 μM oligonucleotide CpG 1668 (sequence 5'-TCCATGA CGTTCCTGATGCT-3'), fully phosphothioated GeneWorks] to promote cell cycling and enable retroviral infection. After overnight culture, cells were washed and resuspended in medium at 0.5 × 10<sup>6</sup> cells/ml, without CpG but containing anti-μ [10 μg/ml AffiniPure F(ab')<sub>2</sub> fragment, goat anti mouse Jackson Laboratories] and/or Baff (250 ng/ml a kind gift from Jürg Tschopp). Cell survival in transduced (GFP<sup>+</sup>) cells was assessed after a further 48 h by flow cytometry, propidium iodide exclusion, and cell counting using internal microbead controls.

Retroviruses expressing transcription factors fused to the human estrogen receptor (hER) dimerization domain were generated by amplification of each factor's ORE, and sequencing each amplified product to ensure it was mutation-free and in frame with the hER. Production of ER fusion proteins of the correct size was confirmed by western blots of infected or transfected cells. Anti-CD40 (clone FGK4.5) was prepared in house and used at 10 μg/ml. β-Estradiol (Sigma) was used at 10 μM, cycloheximide at 50 μM.

## MICE

*Oct2*<sup>-/-</sup>, *Obf1*<sup>-/-</sup>, and *Stat6*<sup>-/-</sup> mice have been described previously (19, 36) *Xbp1*<sup>fl/fl</sup> *Cd19*<sup>Cre/+</sup> mice, where *Xbp1* is conditionally deleted in the B cell lineage, were generated by Hetz et al. (77) and further described in Taubenheim et al. (72). All mice were maintained on a C57BL/6 background, and all experiments conformed

to the relevant regulatory standards of the Animal Ethics Committee of the Walter and Eliza Hall Institute of Medical Research (AEC Projects 2010.010 and 2013.014).

## ANTIBODIES

For westerns, antibodies used were specific for Oct2 (clone 9A2, in house), Obf1 (clone 6F10, in house) Stat6 (S-20), Syk (N-19), and actin (I-19) all from Santa Cruz and Xbp1 (Ab37152-100, Abcam). For flow cytometry, antibodies used were specific for B220 (RA3-6B2), IgD (11-26c.2a), fas/CD95 (Jo2), all from BD Pharmingen Slamf1/CD150 (TC15-12 F12.2, Biolegend), and GL7 (eBioscience). Calcium flux was assessed cytometrically as described (78).

## MICROARRAYS

Illumina Sentrix Mouse v1.1 arrays were probed with RNA prepared from independent B lymphoma cell lines: to identify Oct2-dependent genes, two independent clones of OM1 cells stably transduced with a control vector or one expressing an estradiol-inducible Oct-ER fusion protein were treated *in vitro* for 6 or 48 h with estradiol prior to RNA preparation. For Obf1 targets, two clones of the *Obf1*<sup>-/-</sup> BM1 lymphoma line (78), transduced with vector control or an Obf1-ER expression vector, were induced in a similar manner and RNA prepared and analyzed.

## RNA SEQUENCING

Peripheral B cell populations were sorted from naïve or immunized C57BL/6 mice as described in the legend to **Figure 1**. Two independent biological replicates were prepared for each population, except for spleen and BM ASC. Because of the paucity of ASC in these tissues, cells were pooled from three to four individuals before sorting. Two such pools were processed independently for sequencing. Normalized expression levels are shown in the graphs. As more than 70% of reads in RNA from ASC map to the Ig loci, all Ig reads were excluded from the data, for all populations, before normalization to generate the values shown in **Figure 1A**. Specifically, all reads from the IgH locus on chromosome 12, NC\_000078.6 (positions 113,258,768–116,009,954), all reads from the Igλ locus on chromosome 16, NC\_000082.6 (positions 19,026,858–19,260,844), and all reads from the Igk locus on chromosome 6, NC\_000072.6 (positions 67,555,636–70,726,754) were excluded. For **Figure 2B**, *Oct2*<sup>+/+</sup> and *Oct2*<sup>-/-</sup> B cells were sorted, as in **Figure 2A**, from spleen of two independent mice of each genotype and activated for 48 h before RNA extraction and preparation for sequencing. As these conditions do not induce ASC differentiation, Ig sequence reads were not excluded from analysis of these samples.

## PRIMERS

The starts of *Syk* transcription in primary B cells were determined directly using FirstChoice® RLM-RACE (Ambion) and nested *Syk* gene specific primers specific for the first *Syk* coding exon: 5'-GTAGGTCAGGTGGTTGGCGCTGTCCACAGC-3' and 5'-CCCGCCATGTCTGCACCCCTTCAGAGTTC-3'.

For exon-specific *Syk* qPCR, these primers were used:

5' *Syk* exon 1: 5'-CAGTGACTGCGGCTGAGCGCGGACC-3'  
5' *Syk* exon 2: 5'-CAGCAGGAAACCTCCACTTGCTCTCC-3'

Common 3' *Syk* primer: 5'-CCATGTCTGCACCCCTTCAGAG TTC-3'.

For *Syk* ChIP PCR, these primers were used:

Exon 2: fwd 5'-GCCTAGGCCACGATGGTCAAAGGAGG-3' and rev 5'-GGAGAGCAAGTGGAGGTTTCTGCTG-3'.  
Upstream of exon 2: fwd 5'-CCATTGGTGGGCCCTCAGCTTG GTTC-3' and rev 5'-GACCAGAGAAGAAATGGCCTCAGAAGACAGG-3'.

Chromatin immunoprecipitation was conducted as previously described (29), using high titer, polyclonal anti-Oct2 rabbit serum generated in house.

Primers for other qPCR were:

*Stat6*: fwd 5'-CTGCTGGGCCGAGGCTTCACATTT-3' rev 5'-TCAGGGGCCATTTCCAAGATCATAAGGT-3'  
*CD36*: fwd 5'-GGAGGCATTCTCATGCCAGTCGGAGAC-3' rev 5'-CAAAACTGTCTGTACACAGTGGTGCCTG-3'  
*Slamf1*: fwd 5'-GGGAGCTATCCAGATCACCTG-3' rev 5'-CGTTCTCCTCCACGCTCAC-3'  
*Cd23*: fwd 5'-GACACTGCAATTCAGAATGTCTCTCATG-3' rev 5'-GCTTCTGTTCAGCTTGGAGTTCTTGCAAG-3'  
*Nfil3*: fwd 5'-CTCACGGACCAGGGAGCAGAACCACG-3' rev 5'-CAGGTCTTAAGGACTTCAGCCTCTCATCC-3'  
*Obf1*: fwd 5'-CGGTGTTGACCTATGCTTCTCCACC-3' rev 5'-GAGGGGCGCCTGGTGTCTCGGGACCC-3'  
*Vdr*: fwd 5'-CGCTATGACCTGTGAAGGCTGCAAGGG-3' rev 5'-GCCAATGTCCACGCAGCGTTTGAGCC-3'  
*Xbp1*: fwd 5'-AGCAGCAAGTGGTGGATTG-3' rev 5'-CCAAGCGTGTCTTAACTCCT-3'  
*Hmbs* (for normalization): fwd 5'-GACCTGGTTGTTCACTCC CTGAAG-3' rev 5'-GACAACAGCATCACAAGGGTTTTTC-3'

Chromatin immunoprecipitation PCR and qPCR were performed on triplicate samples in all cases, and the data presented as means  $\pm$  SD. Statistical significance was determined using the unpaired Student's *t*-test.

## IMMUNIZATION

Mice stably reconstituted with *Oct2*<sup>+/+</sup> or *Oct2*<sup>-/-</sup> fetal liver were immunized i.p with  $2 \times 10^9$  sheep red blood cells (Applied Biological Product Management, Australia) in 100  $\mu$ l of PBS and sacrificed after 9 days.

## AUTHOR CONTRIBUTIONS

Lynn Corcoran performed many of the experiments and wrote the manuscript; Dianne Emslie and Tobias Kratina together performed all of the qPCR and protein analyses; Wei Shi performed all bioinformatics analysis for the RNAseq experiments; Susanne Hirsch performed the Slamf1 studies; Nadine Taubenheim contributed to the Xbp1 studies; Stephane Chevrier led most of the flow cytometric analyses.

## ACKNOWLEDGMENTS

We thank Drs. Laurie Glimcher for conditional *Xbp1* mutant mice and Patrick Matthias for *Obf1* mice, Gordon Smythe for bioinformatics assistance and Stephen Nutt for critical comments on the manuscript. Jennifer Vasiliadis and Louise Inglis provided expert animal care. This work was made possible through Victorian State Government Operational Infrastructure Support and Australian Government NHMRC IRIIS and research grants from the NHMRC (#637306 and #575500).

## REFERENCES

1. Staudt LM, Clerc RG, Singh H, Lebowitz JH, Sharp PA, Baltimore D. Cloning of a lymphoid-specific cDNA encoding a protein binding the regulatory octamer DNA motif. *Science* (1988) **241**:577–80. doi:10.1126/science.3399892
2. Herr W, Sturm RA, Clerc RG, Corcoran LM, Baltimore D, Sharp PA, et al. The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2, and *Caenorhabditis elegans* unc-86 gene products. *Genes Dev* (1988) **2**:1513–6. doi:10.1101/gad.2.12a.1513
3. Parslow TG, Blair DL, Murphy WJ, Granner DK. Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. *Proc Natl Acad Sci U S A* (1984) **81**:2650–4. doi:10.1073/pnas.81.9.2650
4. Landolfi NF, Capra JD, Tucker PW. Interaction of cell-type-specific nuclear proteins with immunoglobulin VH promoter region sequences. *Nature* (1986) **323**:548–51. doi:10.1038/323548a0
5. Strubin M, Newell JW, Matthias P. OBF-1, a novel B cell-specific coactivator that stimulates immunoglobulin promoter activity through association with octamer-binding proteins. *Cell* (1995) **80**:497–506. doi:10.1016/0092-8674(95)90500-6
6. Luo Y, Roeder RG. Cloning, functional characterization, and mechanism of action of the B-cell-specific transcriptional coactivator Oca-B. *Mol Cell Biol* (1995) **15**:4115–24.
7. Gstaiger M, Georgiev O, Van Leeuwen H, Van Der Vliet P, Schaffner W. The B cell coactivator Bob1 shows DNA sequence-dependent complex formation with Oct-1/Oct-2 factors, leading to differential promoter activation. *EMBO J* (1996) **15**:2781–90.
8. Tomilin A, Remenyi A, Lins K, Bak H, Leidel S, Vriend G, et al. Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell* (2000) **103**:853–64. doi:10.1016/S0092-8674(00)00189-6
9. Tantin D. Oct transcription factors in development and stem cells: insights and mechanisms. *Development* (2013) **140**:2857–66. doi:10.1242/dev.095927
10. Lins K, Remenyi A, Tomilin A, Massa S, Wilmanns M, Matthias P, et al. OBF1 enhances transcriptional potential of Oct1. *EMBO J* (2003) **22**:2188–98. doi:10.1093/emboj/cdg199
11. Karnowski A, Chevrier S, Belz GT, Mount A, Emslie D, D'Costa K, et al. B and T cells collaborate in antiviral responses via IL-6, IL-21, and transcriptional activator and coactivator, Oct2 and OBF-1. *J Exp Med* (2012) **209**:2049–64. doi:10.1084/jem.20111504
12. Stoykova AS, Sterrer S, Erselius JR, Hatzopoulos AK, Gruss P. Mini-Oct and Oct-2c: two novel, functionally diverse murine Oct-2 gene products are differentially expressed in the CNS. *Neuron* (1992) **8**:541–58. doi:10.1016/0896-6273(92)90282-1
13. Dent CL, Lillycrop KA, Estridge JK, Thomas NS, Latchman DS. The B-cell and neuronal forms of the octamer-binding protein Oct-2 differ in DNA-binding specificity and functional activity. *Mol Cell Biol* (1991) **11**:3925–30.
14. Ninkina NN, Buchman VL, Akopian AN, Lawson SN, Yamamoto M, Campbell E, et al. Nerve growth factor-regulated properties of sensory neurones in Oct-2 null mutant mice. *Brain Res Mol Brain Res* (1995) **33**:233–44. doi:10.1016/0169-328X(95)00128-F
15. Zhou L, Nazarian AA, Xu J, Tantin D, Corcoran LM, Smale ST. An inducible enhancer required for Il12b promoter activity in an insulated chromatin environment. *Mol Cell Biol* (2007) **27**:2698–712. doi:10.1128/MCB.00788-06
16. Cockerill PN, Klinken SP. Octamer-binding proteins in diverse hemopoietic cells. *Mol Cell Biol* (1990) **10**:1293–6.
17. Kamps MP, Corcoran L, Lebowitz JH, Baltimore D. The promoter of the human interleukin-2 gene contains two octamer-binding sites and is partially activated by the expression of Oct-2. *Mol Cell Biol* (1990) **10**:5464–72.

18. Kang SM, Tsang W, Doll S, Scherle P, Ko HS, Tran AC, et al. Induction of the POU domain transcription factor Oct-2 during T-cell activation by cognate antigen. *Mol Cell Biol* (1992) **12**:3149–54.
19. Corcoran LM, Karvelas M, Nossal GJ, Ye ZS, Jacks T, Baltimore D. Oct-2, although not required for early B-cell development, is critical for later B-cell maturation and for postnatal survival. *Genes Dev* (1993) **7**:570–82. doi:10.1101/gad.7.4.570
20. Hatzopoulos AK, Stoykova AS, Erselius JR, Goulding M, Neuman T, Gruss P. Structure and expression of the mouse Oct2a and Oct2b, two differentially spliced products of the same gene. *Development* (1990) **109**:349–62.
21. Annweiler A, Zwilling S, Wirth T. Functional differences between the Oct2 trans-activation domains determine the transactivation potential of individual Oct2 isoforms. *Nucleic Acids Res* (1994) **22**:4250–8. doi:10.1093/nar/22.20.4250
22. Corcoran LM, Koentgen F, Dietrich W, Veale M, Humbert PO. All known in vivo functions of the Oct-2 transcription factor require the C-terminal protein domain. *J Immunol* (2004) **172**:2962–9.
23. Qin XF, Reichlin A, Luo Y, Roeder RG, Nussenzweig MC. Oca-B integrates B cell antigen receptor-, CD40L- and IL 4-mediated signals for the germinal center pathway of B cell development. *EMBO J* (1998) **17**:5066–75. doi:10.1093/emboj/17.17.5066
24. Zwilling S, Dieckmann A, Pfisterer P, Angel P, Wirth T. Inducible expression and phosphorylation of coactivator BOB.1/OBF.1 in T cells. *Science* (1997) **277**:221–5. doi:10.1126/science.277.5323.221
25. Yu X, Wang L, Luo Y, Roeder RG. Identification and characterization of a novel Oca-B isoform. Implications for a role in B cell signaling pathways. *Immunity* (2001) **14**:157–67. doi:10.1016/S1074-7613(09)00090-9
26. Siegel R, Kim U, Patke A, Yu X, Ren X, Tarakhovskiy A, et al. Nontranscriptional regulation of Syk by the coactivator Oca-B is required at multiple stages of B cell development. *Cell* (2006) **125**:761–74. doi:10.1016/j.cell.2006.03.036
27. Humbert PO, Corcoran LM. Oct-2 gene disruption eliminates the peritoneal B-1 lymphocyte lineage and attenuates B-2 cell maturation and function. *J Immunol* (1997) **159**:5273–84.
28. Samardzic T, Gerlach J, Muller K, Marinkovic D, Hess J, Nitschke L, et al. CD22 regulates early B cell development in BOB.1/OBF.1-deficient mice. *Eur J Immunol* (2002) **32**:2481–9. doi:10.1002/1521-4141(200209)32:9<2481::AID-IMMU2481>3.0.CO;2-C
29. Emslie D, D'Costa K, Hasbold J, Metcalf D, Takatsu K, Hodgkin PO, et al. Oct2 enhances antibody-secreting cell differentiation through regulation of IL-5 receptor alpha chain expression on activated B cells. *J Exp Med* (2008) **205**:409–21. doi:10.1084/jem.20072049
30. Hess J, Nielsen PJ, Fischer KD, Bujard H, Wirth T. The B lymphocyte-specific coactivator BOB.1/OBF.1 is required at multiple stages of B-cell development. *Mol Cell Biol* (2001) **21**:1531–9. doi:10.1128/MCB.21.5.1531-1539.2001
31. Jankovic M, Nussenzweig MC. OcaB regulates transitional B cell selection. *Int Immunol* (2003) **15**:1099–104. doi:10.1093/intimm/dxg109
32. Chevrier S, Kratina T, Emslie D, Karnowski A, Corcoran LM. Germinal center-independent, IgM-mediated autoimmunity in sanroque mice lacking OBF1. *Immunol Cell Biol* (2013) **92**(1):12–9. doi:10.1038/icb.2013.71
33. Samardzic T, Marinkovic D, Nielsen PJ, Nitschke L, Wirth T. BOB.1/OBF.1 deficiency affects marginal-zone B-cell compartment. *Mol Cell Biol* (2002) **22**:8320–31. doi:10.1128/MCB.22.23.8320-8331.2002
34. Kim U, Qin XF, Gong S, Stevens S, Luo Y, Nussenzweig M, et al. The B-cell-specific transcription coactivator OCA-B/OBF-1/Bob-1 is essential for normal production of immunoglobulin isotypes. *Nature* (1996) **383**:542–7. doi:10.1038/383542a0
35. Nielsen PJ, Georgiev O, Lorenz B, Schaffner W. B lymphocytes are impaired in mice lacking the transcriptional co-activator Bob1/Oca-B/OBF1. *Eur J Immunol* (1996) **26**:3214–8. doi:10.1002/eji.1830261255
36. Schubart DB, Rolink A, Kosco-Vilbois MH, Botteri F, Matthias P. B-cell-specific coactivator OBF-1/OCA-B/BOB1 required for immune response and germinal centre formation. *Nature* (1996) **383**:538–42. doi:10.1038/383538a0
37. Fehr T, Lopez-Macias C, Odermatt B, Torres RM, Schubart DB, O'Keefe TL, et al. Correlation of anti-viral B cell responses and splenic morphology with expression of B cell-specific molecules. *Int Immunol* (2000) **12**:1275–84. doi:10.1093/intimm/12.9.1275
38. Corcoran LM, Karvelas M. Oct-2 is required early in T cell-independent B cell activation for G1 progression and for proliferation. *Immunity* (1994) **1**:635–45. doi:10.1016/1074-7613(94)90035-3
39. Schubart K, Massa S, Schubart D, Corcoran LM, Rolink AG, Matthias P. B cell development and immunoglobulin gene transcription in the absence of OCT-2 and OBF-1. *Nat Immunol* (2001) **2**:69–74. doi:10.1038/83190
40. Greenwalt DE, Lipsky RH, Ockenhouse CF, Ikeda H, Tandon NN, Jamieson GA. Membrane glycoprotein CD36: a review of its roles in adherence, signal transduction, and transfusion medicine. *Blood* (1992) **80**:1105–15.
41. König H, Pfisterer P, Corcoran LM, Wirth T. Identification of CD36 as the first gene dependent on the B-cell differentiation factor Oct-2. *Genes Dev* (1995) **9**:1598–607. doi:10.1101/gad.9.13.1598
42. Shore P, Dietrich W, Corcoran LM. Oct-2 regulates CD36 gene expression via a consensus octamer, which excludes the co-activator OBF-1. *Nucleic Acids Res* (2002) **30**:1767–73. doi:10.1093/nar/30.8.1767
43. Pfisterer P, Hess J, Wirth T. Identification of target genes of the lymphoid-specific transcription factor Oct2. *Immunobiology* (1997) **198**:217–26. doi:10.1016/S0171-2985(97)80042-3
44. Corcoran L, Vremec D, Febbraio M, Baldwin T, Handman E. Differential regulation of CD36 expression in antigen-presenting cells: Oct-2 dependence in B lymphocytes but not dendritic cells or macrophages. *Int Immunol* (2002) **14**:1099–104. doi:10.1093/intimm/dxf075
45. Hasbold J, Corcoran LM, Tarlinton DM, Tangye SG, Hodgkin PD. Evidence from the generation of immunoglobulin G-secreting cells that stochastic mechanisms regulate lymphocyte differentiation. *Nat Immunol* (2004) **5**:55–63. doi:10.1038/ni1016
46. Nurieva RI, Chung Y, Hwang D, Yang XO, Kang HS, Ma L, et al. Generation of T follicular helper cells is mediated by interleukin-21 but independent of T helper 1, 2, or 17 cell lineages. *Immunity* (2008) **29**:138–49. doi:10.1016/j.immuni.2008.05.009
47. Kallies A, Hasbold J, Tarlinton DM, Dietrich W, Corcoran LM, Hodgkin PD, et al. Plasma cell ontogeny defined by quantitative changes in blimp-1 expression. *J Exp Med* (2004) **200**:967–77. doi:10.1084/jem.20040973
48. Corcoran LM, Tawfils S, Barlow LJ. Generation of B lymphoma cell lines from knockout mice by transformation in vivo with an Emu-myc transgene. *J Immunol Methods* (1999) **228**:131–8. doi:10.1016/S0022-1759(99)00094-0
49. Mackay F, Schneider P, Rennert P, Browning J. BAFF AND APRIL: a tutorial on B cell survival. *Annu Rev Immunol* (2003) **21**:231–64. doi:10.1146/annurev.immunol.21.120601.141152
50. Schwartzberg PL, Mueller KL, Qi H, Cannons JL. SLAM receptors and SAP influence lymphocyte interactions, development and function. *Nat Rev Immunol* (2009) **9**:39–46. doi:10.1038/nri2456
51. Cannons JL, Qi H, Lu KT, Dutta M, Gomez-Rodriguez J, Cheng J, et al. Optimal germinal center responses require a multistage T cell:B cell adhesion process involving integrins, SLAM-associated protein, and CD84. *Immunity* (2010) **32**:253–65. doi:10.1016/j.immuni.2010.01.010
52. Corcoran LM, Hasbold J, Dietrich W, Hawkins E, Kallies A, Nutt SL, et al. Differential requirement for OBF-1 during antibody-secreting cell differentiation. *J Exp Med* (2005) **201**:1385–96. doi:10.1084/jem.20042325
53. Wurster AL, Tanaka T, Grusby MJ. The biology of Stat4 and Stat6. *Oncogene* (2000) **19**:2577–84. doi:10.1038/sj.onc.1203485
54. Tirosch B, Iwakoshi NN, Glimcher LH, Ploegh HL. Rapid turnover of unspliced XBP-1 as a factor that modulates the unfolded protein response. *J Biol Chem* (2006) **281**:5852–60. doi:10.1074/jbc.M509061200
55. Cancro MP. Peripheral B-cell maturation: the intersection of selection and homeostasis. *Immunol Rev* (2004) **197**:89–101. doi:10.1111/j.0105-2896.2004.0099.x
56. Bartholdy B, Du Roure C, Bordon A, Emslie D, Corcoran LM, Matthias P. The Ets factor Spi-B is a direct critical target of the coactivator OBF-1. *Proc Natl Acad Sci U S A* (2006) **103**:11665–70. doi:10.1073/pnas.0509430103
57. Garrett-Sinha LA, Su GH, Rao S, Kabak S, Hao Z, Clark MR, et al. PU.1 and Spi-B are required for normal B cell receptor-mediated signal transduction. *Immunity* (1999) **10**:399–408. doi:10.1016/S1074-7613(00)80040-0
58. Hu CJ, Rao S, Ramirez-Bergeron DL, Garrett-Sinha LA, Gerondakis S, Clark MR, et al. PU.1/Spi-B regulation of c-rel is essential for mature B cell survival. *Immunity* (2001) **15**:545–55. doi:10.1016/S1074-7613(01)00219-9
59. Dienz O, Eaton SM, Bond JP, Neveu W, Moquin D, Noubade R, et al. The induction of antibody production by IL-6 is indirectly mediated by IL-21 produced by CD4+ T cells. *J Exp Med* (2009) **206**:69–78. doi:10.1084/jem.20081571
60. Eto D, Lao C, Ditoro D, Barnett B, Escobar TC, Kageyama R, et al. IL-21 and IL-6 are critical for different aspects of B cell immunity and redundantly induce

- optimal follicular helper CD4 T cell (T<sub>fh</sub>) differentiation. *PLoS One* (2011) **6**:e17739. doi:10.1371/journal.pone.0017739
61. Haynes NM, Allen CD, Lesley R, Ansel KM, Killeen N, Cyster JG. Role of CXCR5 and CCR7 in follicular Th cell positioning and appearance of a programmed cell death gene-1-high germinal center-associated subpopulation. *J Immunol* (2007) **179**:5099–108.
  62. Glatman Zaretsky A, Taylor JJ, King IL, Marshall FA, Mohrs M, Pearce EJ. T follicular helper cells differentiate from Th2 cells in response to helminth antigens. *J Exp Med* (2009) **206**:991–9. doi:10.1084/jem.20090303
  63. Deenick EK, Chan A, Ma CS, Gatto D, Schwartzberg PL, Brink R, et al. Follicular helper T cell differentiation requires continuous antigen presentation that is independent of unique B cell signaling. *Immunity* (2010) **33**:241–53. doi:10.1016/j.immuni.2010.07.015
  64. Baumjohann D, Okada T, Ansel KM. Cutting edge: distinct waves of BCL6 expression during T follicular helper cell development. *J Immunol* (2011) **187**:2089–92. doi:10.4049/jimmunol.1101393
  65. Baumjohann D, Preite S, Reboldi A, Ronchi F, Ansel KM, Lanzavecchia A, et al. Persistent antigen and germinal center B cells sustain T follicular helper cell responses and phenotype. *Immunity* (2013) **38**:596–605. doi:10.1016/j.immuni.2012.11.020
  66. Goenka R, Barnett LG, Silver JS, O'Neill PJ, Hunter CA, Cancro MP, et al. Cutting edge: dendritic cell-restricted antigen presentation initiates the follicular helper T cell program but cannot complete ultimate effector differentiation. *J Immunol* (2011) **187**:1091–5. doi:10.4049/jimmunol.1100853
  67. Paul WE, Ohara J. B-cell stimulatory factor-1/interleukin 4. *Annu Rev Immunol* (1987) **5**:429–59. doi:10.1146/annurev.immunol.5.1.429
  68. Casellas R, Jankovic M, Meyer G, Gazumyan A, Luo Y, Roeder R, et al. OcaB is required for normal transcription and V(D)J recombination of a subset of immunoglobulin kappa genes. *Cell* (2002) **110**:575–85. doi:10.1016/S0092-8674(02)00911-X
  69. Reimold AM, Iwakoshi NN, Manis J, Vallabhajosyula P, Szomolanyi-Tsuda E, Gravalles EM, et al. Plasma cell differentiation requires the transcription factor XBP-1. *Nature* (2001) **412**:300–7. doi:10.1038/35085509
  70. Iwakoshi NN, Lee AH, Vallabhajosyula P, Otipoby KL, Rajewsky K, Glimcher LH. Plasma cell differentiation and the unfolded protein response intersect at the transcription factor XBP-1. *Nat Immunol* (2003) **4**:321–9. doi:10.1038/ni907
  71. Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, Lee AH, Qian SB, Zhao H, et al. XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* (2004) **21**:81–93. doi:10.1016/j.immuni.2004.06.010
  72. Taubenheim N, Tarlinton DM, Crawford S, Corcoran LM, Hodgkin PD, Nutt SL. High rate of antibody secretion is not integral to plasma cell differentiation as revealed by XBP-1 deficiency. *J Immunol* (2012) **189**:3328–38. doi:10.4049/jimmunol.1201042
  73. Shen Y, Hendershot LM. Identification of ERdj3 and OBF-1/BOB-1/OCA-B as direct targets of XBP-1 during plasma cell differentiation. *J Immunol* (2007) **179**:2969–78.
  74. Nutt SL, Taubenheim N, Hasbold J, Corcoran LM, Hodgkin PD. The genetic network controlling plasma cell differentiation. *Semin Immunol* (2011) **23**:341–9. doi:10.1016/j.smim.2011.08.010
  75. Boyd AW, Goding JW, Schrader JW. The regulation of growth and differentiation of a murine B cell lymphoma. I. Lipopolysaccharide-induced differentiation. *J Immunol* (1981) **126**:2461–5.
  76. Krebs DL, Yang Y, Dang M, Haussmann J, Gold MR. Rapid and efficient retrovirus-mediated gene transfer into B cell lines. *Methods Cell Sci* (1999) **21**:57–68. doi:10.1023/A:1009843325770
  77. Hetz C, Lee AH, Gonzalez-Romero D, Thielen P, Castilla J, Soto C, et al. Unfolded protein response transcription factor XBP-1 does not influence prion replication or pathogenesis. *Proc Natl Acad Sci U S A* (2008) **105**:757–62. doi:10.1073/pnas.0711094105
  78. Veale MF, Dietrich WM, Corcoran LM. Knockout B lymphoma cell lines as biochemical tools to explore multiple signalling pathways. *Immunol Cell Biol* (2003) **81**:297–304. doi:10.1046/j.0818-9641.2003.01172.x

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 December 2013; accepted: 03 March 2014; published online: 20 March 2014.

Citation: Corcoran L, Emslie D, Kratina T, Shi W, Hirsch S, Taubenheim N and Chevrier S (2014) Oct2 and Obf1 as facilitators of B:T cell collaboration during a humoral immune response. *Front. Immunol.* 5:108. doi: 10.3389/fimmu.2014.00108  
This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Corcoran, Emslie, Kratina, Shi, Hirsch, Taubenheim and Chevrier. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Regulation of the NF- $\kappa$ B-mediated transcription of inflammatory genes

Dev Bhatt and Sankar Ghosh\*

Department of Microbiology and Immunology, College of Physicians and Surgeons, Columbia University, New York, NY, USA

**Edited by:**

Ananda L. Roy, Tufts University School of Medicine, USA

**Reviewed by:**

John D. Colgan, University of Iowa, USA

Stephen Smale, University of California Los Angeles, USA

**\*Correspondence:**

Sankar Ghosh, Department of Microbiology and Immunology, College of Physicians and Surgeons, Columbia University, HHSC 1210, 701 West 168th Street, New York, NY 10032, USA  
e-mail: sg2715@columbia.edu

The NF- $\kappa$ B family of transcription factors plays a central role in the inducible expression of inflammatory genes during the immune response, and the proper regulation of these genes is a critical factor in the maintenance of immune homeostasis. The chromatin environment at stimulus-responsive NF- $\kappa$ B sites is a major determinant in transcription factor binding, and dynamic alteration of the chromatin state to facilitate transcription factor binding is a key regulatory mechanism. NF- $\kappa$ B is in turn able to influence the chromatin state through a variety of mechanisms, including the recruitment of chromatin modifying co-activator complexes such as p300, the competitive eviction of negative chromatin modifications, and the recruitment of components of the general transcriptional machinery. Frequently, the selective interaction with these co-activators is dependent on specific post-translational modification of NF- $\kappa$ B subunits. Finally, the mechanisms of inducible NF- $\kappa$ B activity in different immune cell types seem to be largely conserved. The diversity of cell-specific NF- $\kappa$ B-mediated transcriptional programs is established at the chromatin level during cell differentiation by lineage-defining transcription factors. These factors generate and maintain a cell-specific chromatin landscape that is accessible to NF- $\kappa$ B, thus restricting the inducible transcriptional response to a cell-appropriate output.

**Keywords:** transcription, chromatin, gene expression, NF-kappaB, signaling, transcription factor

## INTRODUCTION

Upon pathogen detection, the innate immune system must be able to mount a robust and rapid response, but equally important is the need to rein in the cytotoxic effects of the inflammatory response. Therefore, modulation of pro-inflammatory gene expression is of fundamental importance in the maintenance of immune homeostasis. Pro-inflammatory genes are maintained in a silent, yet poised, state that can be rapidly induced in response to different stimuli, and this characteristic pattern is achieved through the action of two elements: the activation of inducible transcription factors and the modulation of the chromatin environment at gene regulatory elements.

Multiple signaling pathways are activated in response to immune and inflammatory stimuli, resulting in the activation of many different transcription factors. The transcription factors induced upon stimulation must interact with *cis*-regulatory elements of target genes to facilitate recruitment of the general transcriptional machinery. The chromatin state at these *cis*-elements plays a critical role in modulating the activity of transcription factors, mainly by functioning as a steric barrier to DNA binding and as a post-translational regulatory platform that influences the recruitment of transcriptional cofactors. This review will focus on the interplay of the archetypal inducible transcription factor NF- $\kappa$ B with the chromatin environment, and discuss how the chromatin presents a selective regulatory barrier to NF- $\kappa$ B activity and how NF- $\kappa$ B alters the chromatin environment to induce transcription of inflammatory genes.

The NF- $\kappa$ B family of transcription factors is conserved through metazoan organisms. These transcription factors are characterized

by a unique DNA-binding motif known as the Rel homology domain (RHD). In mammals, there are five RHD containing proteins: p65 (RelA), c-Rel, RelB, p100/p50, and p105/p52. Each protein is capable of forming homodimers and heterodimers, with 15 dimer combinations possible (1). In unstimulated cells, NF- $\kappa$ B is sequestered in the cytoplasm by the inhibitor of NF- $\kappa$ B family (I $\kappa$ B) via their ankyrin repeat domains. Upon stimulation, the I $\kappa$ B kinase (IKK) complex is activated and phosphorylates serine residues on I $\kappa$ B molecules, targeting them for ubiquitination by the SCF E3 ligase and subsequent degradation by the proteasome. Degradation of I $\kappa$ B releases NF- $\kappa$ B, which translocates into the nucleus to initiate a transcriptional response (2).

## BASIC TRANSCRIPTIONAL MECHANISMS

Transcription can fundamentally be understood as the interaction of *cis*- and *trans*-acting factors within the nucleus that orchestrate the expression of a particular gene. The *cis*-acting elements are defined as non-coding DNA elements located on the same chromosome as the protein-coding locus. Two critical *cis*-elements are the promoter and the enhancer. *Trans*-acting factors include sequence-specific transcription factors (such as NF- $\kappa$ B and IRF), chromatin modifying complexes (such as histone acetyltransferases and chromatin remodeling complexes), the mediator complex, and finally general transcriptional factors (GTFs), including RNA polymerase II (3–5).

Transcription initiates from the *cis*-regulatory element known as the promoter, which can be divided into two portions, the core promoter and the proximal promoter. The core promoter contains regulatory elements that bind Pol II and the general transcriptional



machinery, extending from  $-35$  to  $+35$  relative to the transcriptional start site. The distal promoter extends  $\sim 300$  bp upstream from the core promoter and serves as the binding site for sequence-specific transcription factors. It is thought that proximal promoter binding transcription factors coordinate with distal enhancer-binding transcription factors to recruit the mediator complex and the general transcriptional machinery.

A number of conserved elements comprise the core promoter. This includes, but is not limited to, the BRE (TFIIB recognition element), TATA box, initiator, and downstream promoter element (DPE). However, not all of these elements are absolutely required for a promoter to be functional. In addition to these discrete elements, vertebrate promoters are also distinguished by their CpG dinucleotide content, with high CpG content promoters commonly classified as CpG-island promoters (CGI) (6–8). However, all these regulatory elements function within the context of chromatin.

Eukaryotic DNA is highly condensed, a necessary strategy for the compaction of large genomes into a relatively small nuclear volume. This is achieved by the wrapping of DNA around histone proteins. The resultant DNA:histone complex is referred to as the nucleosome, which is the core repeating unit of the chromatin. The degree of compaction of chromatin fibers plays an important role in accessibility of DNA-binding proteins for their cognate binding sites. Furthermore, elongating polymerases encounter nucleosomal barriers during their progression through the locus. Thus, chromatin presents additional layers of complexity to the mechanism of transcription (9).

There are four major histone proteins, H2A, H2b, H3, and H4. Two copies of each protein form the histone octamer, and 147 bp of DNA wraps around the octamer to form the nucleosome. The interaction of DNA and the octamer is highly stable. High-resolution crystal structures have shown that there are over 100 points of contact between the octamer and DNA, and that the DNA is stabilized by arginine residues of the octamer directly contacting the minor groove of DNA (10).

### HISTONES AND CHROMATIN MODIFICATION DOMAINS AT REGULATORY ELEMENTS

The tails of histone proteins are subject to a broad range of post-translational modifications that influence all aspects of chromatin biology. Histone modification is thought to have two major purposes: (1) the alteration of net charge of the tail, which has an effect on tail-DNA interactions and inter-histone interactions; (2) the generation of recognition sites for activating or repressive factors (9).

Lysine acetylation and methylation have been the most well-studied histone modifications that regulate transcription. Addition of acetyl residues is accomplished by histone acetyltransferase complexes (HATs), and removal of these marks is catalyzed by histone deacetylases (HDACs). HATs, such as the p300/CBP complex, are generally non-specific and are able to target multiple lysine residues on various histone proteins. Lysine acetylation is largely associated with active transcription. Acetylation of H4K16 has been shown to have a significant impact on the compaction of nucleosomal arrays *in vitro*, and is a general mark of euchromatin *in vivo*. The specialized domain that recognizes acetylated lysines

is known as the bromodomain. Proteins that have bromodomains include HATs and chromatin remodeling complexes such as RSC and SWI/SNF (11, 12).

Lysine methylation can act as an activating or repressive mark, depending on the specific residue modified. For example, trimethylation of H3K4 is strongly associated with active transcription, and is found at active promoters, while methylation of H3K9 and H3K27 is associated with transcriptional repression. In contrast to acetyltransferases, methyltransferase have very restricted substrate specificities, in keeping with the more specialized and context-dependent roles that methylated lysines have. Furthermore, methylated lysine residues have relatively restricted distributions across gene loci; for example, H3K4me3 is enriched at promoters of active genes, while H3K4me1 has been more recently recognized as an enhancer-specific mark, and finally H3K36me3 and H3K79me are found within gene bodies (3, 11).

### SETTING THE STAGE: DEVELOPMENTAL ESTABLISHMENT OF AN NF- $\kappa$ B-RESPONSIVE GENE PROGRAM

NF- $\kappa$ B dimers are broadly distributed, especially in cells of the immune system, and are activated in response to a variety of cell-specific receptor stimuli (e.g., TCR in T-cells, pattern recognition receptors in myeloid cells). In response to these stimuli however, NF- $\kappa$ B is able to induce the expression of a diverse range of genes tailored to a specific cellular response. Such cell-type specific response is established during development and is broadly speaking, the result of lineage-defining transcription factors binding to enhancer elements at an early developmental stage. In macrophages for example, the major lineage-determining transcription factors PU.1 and C/EBP and AP-1 families bind cognate enhancer elements and establish an active chromatin state. This renders the enhancer and promoter elements accessible to additional factors, resulting in a cell-specific array of binding sites in genes that have been epigenetically primed for binding to activated NF- $\kappa$ B (13, 14). The establishment of an inducible epigenetic landscape is not the only mechanism by which NF- $\kappa$ B binding is modulated. Many loci require the synergistic activity of multiple signal-dependent transcription factors, as well as further modulation of the chromatin structure at enhancers and promoter elements.

### REGULATION BEFORE THE SIGNAL: CHROMATIN AS A DETERMINANT FOR NF- $\kappa$ B BINDING ACTIVITY

The concerted activity of lineage-defining transcription factors and inducible transcription factors are required for the proper regulation of inflammatory genes expression. Along those lines, the synergistic activity of multiple inducible transcription factors is frequently necessary to overcome the inhibitory chromatin state at inflammatory genes. A classic example of transcription factor synergism is the induction of the human interferon- $\beta$  (*IFNB*) gene in response to viral infection. Stimulus-dependent expression of this gene requires the cooperative binding of three transcription factors: NF- $\kappa$ B, IRF3/IRF7, and ATF-2/c-JUN. NF- $\kappa$ B initially binds to the conserved PRDII element in the promoter. This in turn facilitates the recruitment of IRF and ATF-2/c-Jun. Once properly assembled at the promoter, these transcription factors serve as a platform for the sequential recruitment of the PCAF chromatin

modifying complex, the p300/CBP acetyltransferase, and subsequently the SWI/SNF chromatin remodeling complexes. SWI/SNF remodels the downstream nucleosome that encompasses the TATA box, thus allowing TBP binding and subsequent pre-initiation complex assembly (15–17). The regulation of the IFN $\beta$  enhanceosome is an important illustration of transcriptional regulation through the combinatorial control of multiple transcriptional factors, and how only the concerted effort of these factors can resolve the presence of a chromatin barrier.

Studies on the regulation of expression of the cytokine gene *Il12b* have shown that a non-permissive chromatin configuration functions as regulatory barrier that must be inducibly resolved for proper gene expression. Early studies on *Il12b* expression focused on identifying critical transcription factor binding sites by systematic mutation of promoter-reporter constructs. These studies showed that NF- $\kappa$ B, specifically the c-Rel subunit, in conjunction with C/EBP, AP-1, and NFAT transcription factors, are required for *Il12b* activation (18–21). Nucleosomal mapping of endogenous *Il12b* promoter revealed that the critical transcription factor binding sites were occupied by a positioned nucleosome extending from –30 to –175 upstream of the transcriptional initiation site, and that this nucleosome was selectively remodeled upon LPS stimulation. Furthermore, this promoter remodeling event was dependent on *de novo* protein synthesis, but was independent of c-Rel binding, as evidenced by promoter remodeling occurring in c-Rel deficient cells that lacked the ability to express *Il12b* (22–24). Thus, in contrast to the enhanceosome, which required NF- $\kappa$ B in a synergistic complex to recruit chromatin remodeling complexes, *Il12b* operates under a slightly different mode of regulation in which remodeling is required prior to DNA binding. Moreover, since NF- $\kappa$ B genes are induced with variable kinetics, there would appear to be multiple chromatin-based regulatory strategies governing the binding of NF- $\kappa$ B.

To this end, Sacconi and colleagues made the important observation that although NF- $\kappa$ B RelA subunits entered the nucleus rapidly upon stimulation, they only bound to a subset of genes initially (25). Other genes were bound at later time points, suggesting that their binding sites were inaccessible, consistent with the observations at the *Il12b* promoter. Inspection of the chromatin modifications at early versus late bound genes showed a consistent pattern: early recruiting genes had higher levels of histone acetylation while late bound genes had low basal acetylation levels that increased upon stimulation (25). From these data, a model emerged in which NF- $\kappa$ B binding was inhibited at certain promoters during the early phase of stimulation by inaccessible chromatin structure, and NF- $\kappa$ B could only bind these late gene promoters after critical chromatin remodeling events had occurred.

Implicit in this model is the idea that NF- $\kappa$ B itself lacks the ability to bind a chromatinized template and requires the binding of additional factors capable of recruiting chromatin remodeling complexes. This hypothesis is supported by structural studies of NF- $\kappa$ B, which showed its precise binding to a naked DNA template (26–28). It has been reported that NF- $\kappa$ B p50-homodimers can indeed bind nucleosomal  $\kappa$ B sites *in vitro*, albeit at a reduced efficiency than the naked template (29, 30). Furthermore, the positioning of the binding site within the nucleosome strongly influenced the binding affinity of p50, with binding sites near the

edge of the nucleosome being highly favored. The *in vitro* nucleosome binding was also at least partially dependent on remodeling complex activity or partial disassembly of the histone octamer (31). It remains to be seen whether this nucleosome binding activity is specific to p50-homodimers and its specific cognate site, or whether it is a common feature shared by a broad variety of NF- $\kappa$ B dimers and binding sites. Interestingly, p50 dimer complexes have been observed in the nucleus of unstimulated cells, which are displaced by activating dimers (RelA or c-Rel species) in stimulated cells presumably after chromatin remodeling has occurred (32, 33). Given that p50 lacks a trans-activation domain, and has been shown to associate with deacetylase complexes, the latent p50 binding to a more compact chromatin template may be part of a regulatory strategy keeping genes silent under resting conditions by maintaining a repressive chromatin environment (32).

Although it is clear that differential chromatin states influence the NF- $\kappa$ B response, the mechanisms that contribute to inducible chromatin remodeling remained unclear. To test whether specific chromatin remodeling complexes were required for expression of NF- $\kappa$ B dependent genes, Ramirez-Carrozzi and colleagues targeted the SWI/SNF chromatin remodeling complex by retroviral knockdown of its ATPase subunits, Brg and Brm (34). This protein complex was a likely candidate for chromatin remodeling as it had been shown to be strongly associated with gene activation in various contexts. In the targeted cells, the chromatin remodeling of *Il12b* was strongly inhibited, concurrent with a loss in expression of *Il12b*. The knockdown, however, had no effect on the inducible expression of another gene, the rapidly expressed chemokine, *Cxcl2*. Further comparison of the expression patterns of these two genes revealed that (1) *Cxcl2* was strongly induced upon 30 min of LPS treatment while *Il12b* was most strongly induced only after 2 h of treatment and (2) expression of *Cxcl2* was not inhibited by protein synthesis inhibitors, indicating that it is a primary response gene. Based on such criteria including induction kinetics, protein synthesis requirement, and SWI/SNF dependence, inducible genes could be partitioned into three classes: early primary, late primary, and secondary. The early primary response class was induced rapidly and did not require either Brg1/Brm or new protein synthesis and included genes such as *Cxcl2*, *Tnf*, and *Ptgs2*. However, a subset of primary response genes (which included *Ccl5*, *Saa3*, and *Ifnb1*) did require Brg1/Brm for activation and were induced with delayed kinetics relative to remodeling-independent primary response genes. Finally, secondary response class (which included *Il12b*, *Nos2*, and *Il6*) required both Brg1/Brm and new protein synthesis. ChIP analysis of Brg showed that the remodeling complex was associated in an inducible fashion at remodeling-dependent genes. Furthermore, nuclease sensitivity analysis at representative genes from each class showed a pattern consistent with their dependence on Brg1/Brm. Namely, inducible promoter accessibility was seen at late primary and secondary genes, and this sensitivity was lost during knockdown of Brg1/Brm. In contrast, early primary genes had much more accessible promoters and this accessibility was unchanged by Brg1/Brm knockdown (34, 35).

This classification of inflammatory genes was further expanded by the discovery that many remodeling-independent genes were distinguished by the presence of a CpG island (CGIs) (35, 36). CGIs are prevalent in the promoters of ~70% of protein-coding gene

promoters, including the promoters of ubiquitously expressed housekeeping genes (37). The presence of CpG-island promoters is highly correlated with an accessible chromatin configuration, namely high levels of H3K4me3 marks, as well as pre-association of Pol II in the basal state. The contribution CGIs to this favorable chromatin state is thought to emerge from multiple possible mechanisms. First, the recruitment of chromatin modifying complexes that deposit H3K4me3 marks and display specificity for CGIs via CXXC domain containing proteins (38, 39). Secondly, CGI sequences may be inherently unfavorable to the formation of stable, repressive nucleosome formation at these promoters (35, 40). Finally, many transcription factors, such as Sp1, have binding affinity for GC-rich sequences, and may contribute to the maintenance of an open chromatin configuration of CGI promoters. In contrast, non-CpG-island genes were largely remodeling dependent, and many such genes have been shown to be dependent on the IRF transcription factors, which are competent for recruiting SWI/SNF complexes.

These studies have generated a regulatory framework that integrates the basal chromatin state with the kinetics of transcriptional induction and selective requirement for chromatin remodeling. Within this model, the kinetics of a particular gene's expression correlated with the basal chromatin state and the synthesis and induction of transcription factors (such as IRF) capable of promoting the remodeling of nucleosomal barriers to NF- $\kappa$ B dimer recruitment and transcriptional activation. NF- $\kappa$ B dimers may still be involved in recruitment of remodeling complexes but are most likely to do so in cooperation with other transcription factors (Figure 1).

Subsequent to binding, NF- $\kappa$ B plays a major, yet heterogeneous role, at the chromatin level. The many different interactions between NF- $\kappa$ B and its various transcriptional cofactors have not been extensively defined, and the molecular order of many of these various interactions remain unclear. Nevertheless, it is accepted that NF- $\kappa$ B regulates the expression of its target genes through

a broad range of chromatin-mediated mechanisms. These roles can be generalized into two broadly defined activities: recruitment of positive cofactors/marks and the removal of negative regulators/marks, as discussed below.

## NF- $\kappa$ B IN THE NUCLEUS: A VARIETY OF ROLES AND A VARIETY OF PARTNERS

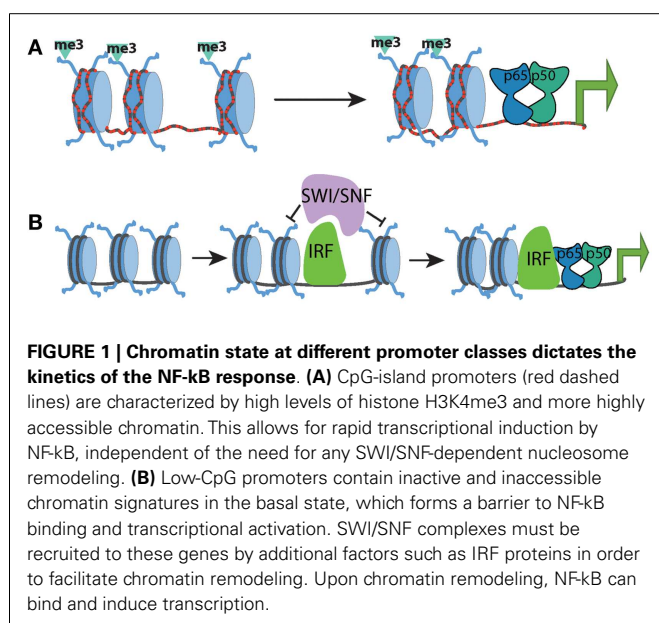
### RECRUITMENT OF CO-ACTIVATORS

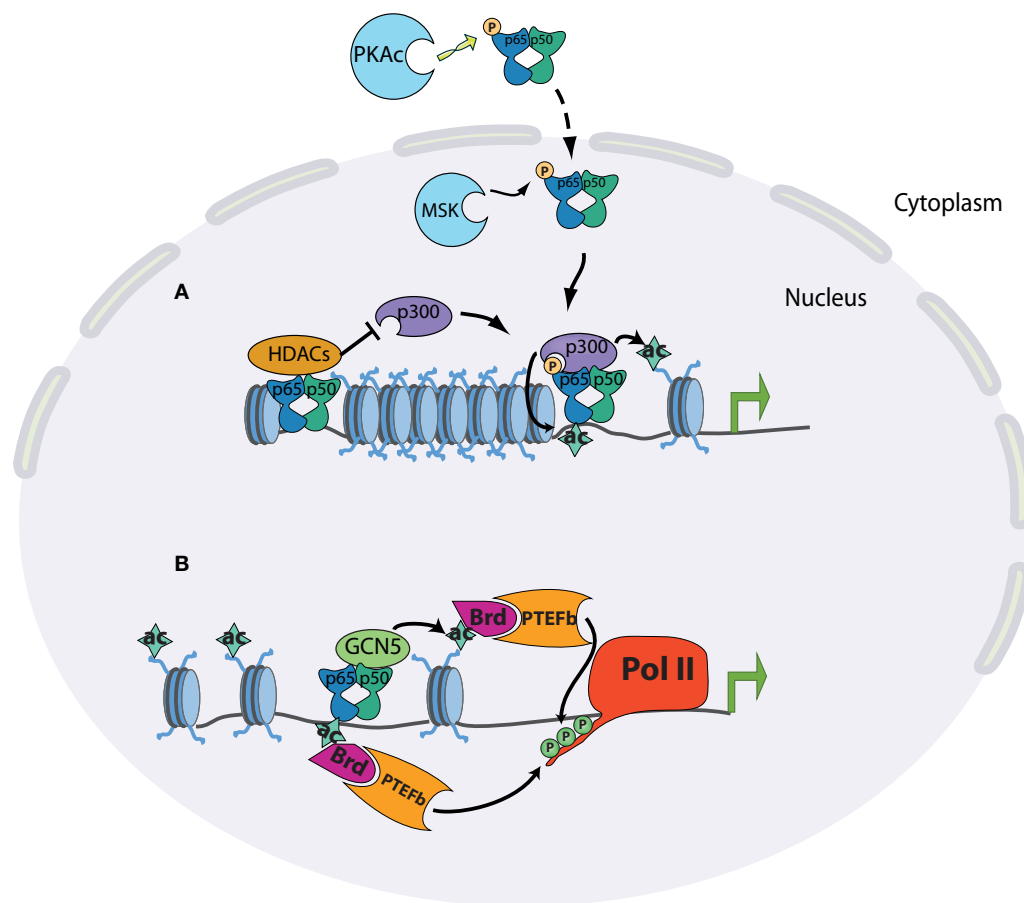
NF- $\kappa$ B is capable of interacting with many different transcriptional cofactors via its trans-activation domain, including chromatin modifying complexes and general transcription factors (41). The purpose of these chromatin modifications is to serve as a recruitment platform for additional activators, such as remodeling complexes, enabling recruitment of the transcriptional machinery to the promoter, followed by initiation of transcription. Although many of these cofactors are essential for the expression of NF- $\kappa$ B-induced genes, the mechanisms of how NF- $\kappa$ B can selectively regulate or recruit them remains unclear.

One of the most well-characterized cofactors of NF- $\kappa$ B, specifically the RelA subunit, is the p300/CBP histone acetyltransferase complex. The interaction between these factors is interesting in that it requires additional modification of RelA in order to occur. Post-translational modifications of RelA have been shown to play an important role in modulating the activity of NF- $\kappa$ B. Although the mechanistic importance of many of these modifications remains unclear, studies of specific residues have shown that they may regulate the recruitment of transcriptional cofactors such as histone acetyltransferases (HATs) or histone deacetylases (HDACs). The p300/CBP complex is an important example of a chromatin modifying complex that directly interacts with NF- $\kappa$ B. Strikingly this interaction is regulated at the post-translational level, being dependent on the specific phosphorylation of the RelA protein.

The Ser276 residue of RelA can be phosphorylated, has been shown to be targeted by the catalytic subunit of protein kinase A (PKA), mitogen- and stress-activated protein kinase-1 (MSK1), and Pim1 kinase (42–45). The importance of this phosphorylated serine was shown to be threefold: it moderately enhanced DNA-binding activity of RelA; it caused a conformational shift that allowed CBP/p300 binding in place of the latent interaction of RelA with HDAC complexes; RelA deficient cells reconstituted with S276A mutants were severely impaired in NF- $\kappa$ B dependent gene expression (32, 45–48) (Figure 2). Most convincingly, the absolute necessity of this phosphorylation event was demonstrated by the generation of a knock-in mouse containing an alanine substitution at RelA S276 residue. Mice homozygous for S276A point mutants were embryonic lethal between E11 and E16 and exhibited severe developmental defects, most notably in eye development, and embryonic fibroblasts showed a defect in the activation of selective subsets of pro-inflammatory genes (47).

The interplay between p300/CBP and RelA illustrates critical aspects of NF- $\kappa$ B biology. In order for NF- $\kappa$ B to fully function, it must be activated by canonical NF- $\kappa$ B signaling, and then modified by a distinct signaling pathway. p300/CBP has also been shown to acetylate RelA itself, notably on lysine 310. As with phosphorylation of Ser276, differential modification of Lys310 plays a major role in dictating RelA cofactor specificity. Acetylation of the residue





**FIGURE 2 | Chromatin state at different promoter classes dictates the kinetics of the NF- $\kappa$ B response. (A)** Release of NF- $\kappa$ B dimers through appropriate stimulation is linked to phosphorylation of p65 at serine 276 in the cytoplasm by PKAc or in the nucleus by MSK1/2. Phosphorylated p65 preferentially interacts with CBP/p300 co-activator complexes, displacing its latent binding to HDAC complexes. The specific interaction of p65 and p300

results in the acetylation of p65 itself and of surrounding histones, and subsequent transcriptional induction. **(B)** The acetylation of p65 on K310 promotes the recruitment of the positive elongation factors Brd4 and P-TEFb, which can phosphorylate the polymerase C-terminus. p65 can also recruit GCN5 acetyltransferase complexes, leading to promoter acetylation, Brd4-P-TEFb recruitment, and polymerase elongation.

promotes an interaction with the histone acetyltransferase Tip60 and the transcriptional elongation factors PTEF-B (49–51). Conversely, methylation of Lys310 by the SET6 proteins facilitates an interaction with the histone methyltransferase GLP/G9a, leading to downregulation of transcription. The selective modulation of these two residues highlights the functional diversity of NF- $\kappa$ B in the control of the inflammatory response. Many other post-translational modifications have been shown to occur on RelA, with a corresponding enhancement or attenuation of transcription of selective inflammatory genes (52, 53).

#### DE-REPRESSION: REMOVAL OF REPRESSIVE MARKS AND COMPLEXES

In addition to the inducible recruitment of activating marks, another regulatory mechanism is to maintain a basal repressive state at gene promoters. These repressive marks are subsequently removed upon stimulation. This phenomenon can occur at the level of basal chromatin marks, or the binding of repressive complexes. For example, some inducible gene promoters have high levels of the H3K9 dimethyl modification, a repressive

mark associated with transcriptional silencing. Upon stimulation, this mark is removed by the Aof1 histone demethylase, which is recruited by initially bound c-Rel dimers (54, 55). Similarly, subsets of inflammatory gene promoters are marked by the repressive trimethyl H3K27 deposited by the Ezh2 methyltransferase, and the inducible removal of this mark by the Jmjd3 demethylase is required for their expression (56). The dynamic regulation of this mark has proved to be of interest, due to the development of a Jmjd3-specific pharmacological inhibitor that can attenuate inducible pro-inflammatory gene expression (57).

The stimulus-dependent eviction of basally bound co-repressor complexes at promoters is another chromatin-based strategy for maintaining tight regulation on the expression of NF- $\kappa$ B dependent genes. Nuclear receptor corepressor (NCoR) and the closely related protein silencing mediator of retinoid and thyroid receptors (SMRT) are multiprotein co-repressor complexes containing histone deacetylases (HDACs), specifically HDAC3. These HDACs function to keep H3K9/14 acetylation levels low in the resting cell. Upon stimulation, this repressive state is relieved, allowing

for the accumulation of activating marks and recruitment of the transcriptional machinery. Similarly, NCoR/SMRT complexes are also displaced, promoting NF- $\kappa$ B binding and recruitment of co-activator complexes. It should be noted that both NCoR and SMRT complexes function in non-redundant fashion, and are recruited to regulatory elements by different sequence-specific transcription factors: JUN homodimers and p50 dimers in the case of NCoR, and the ETS-related protein, translocation-ETS-leukemia (TEL) in the case of SMRT (36, 58, 59). In addition to HDAC-dependent removal of activating marks, NCoR complexes can maintain a silent state via the deposition of the repressive H3K20me3 mark. Subsequently upon stimulation, these marks are removed by the NF- $\kappa$ B dependent recruitment of histone demethylase (60).

Because these repressive complexes restrain the expression of inflammatory gene program in the basal state, one would predict that their targeted disruption would lead to hyperacetylation and therefore enhanced expression of immune genes. Surprisingly however this is not the case. Knockout of NCoR in macrophages led to an anti-inflammatory phenotype, contrary to the expectation that removal of a basal repressor would lead to the hyper-responsiveness of pro-inflammatory genes. Although NF- $\kappa$ B binding activity appeared to be unaffected in these cells, the lack of the basal repressive complex compromised the stimulus-dependent deposition of H3K4me2, a mark associated with productive transcription. The mechanism contributing to this phenotype was attributed to the de-repression of metabolic pathways that subsequently inhibit chromatin modifying complexes required for induction of pro-inflammatory genes (61). Along similar lines, HDAC3-deficient macrophages were defective in their ability to activate a broad range of pro-inflammatory genes. Although the knockout did result in hyperacetylation across immune genes, there was a specific defect in the expression of IFN $\beta$ , which compromised overall STAT protein signaling and the secondary gene response to TLR4 stimulation (62).

#### NF- $\kappa$ B-MEDIATED REGULATION OF POL II ELONGATION

After polymerase recruitment to the promoter, sequential modifications of Pol II are required for productive transcription to occur. The initial stage of transcriptional elongation requires local unwinding of the promoter DNA. Characteristic of this stage is the phosphorylation of serine 5 of the C-terminal heptad repeat by the Cdk7 subunit of TFIIF, or the CDK8 subunit of the mediator complex. Ser5 phosphorylation serves as a platform to recruit RNA capping enzymes to the nascent transcript (63, 64). A second pause occurs ~40–50 bp downstream of the transcription initiation site, and de-repression of this block requires the kinase activity of P-TEFb. This final de-repression commits the polymerase to the elongation phase of transcription (65, 66). During an inflammatory response, there is evidence for stimulus-dependent unpausing being the major rate-limiting step in the expression of rapidly induced genes (36, 67, 68). Along with the p300/CBP complex, NF- $\kappa$ B has been shown to recruit GCN5 acetyltransferase complexes, which primarily modify H4K5/K8/K12 lysines. These residues have been shown to be deposited in response to NF- $\kappa$ B binding. The accumulation of acetylated H4 in stimulated cells allows binding of the bromodomain-containing protein BRD4,

which then plays a positive role in transcription by recruiting the elongation factor P-TEFb (Figure 2) (36).

The signal-dependent elongation by Pol II is most prevalent at rapidly induced CGI containing promoters, where the permissive chromatin structure allows for a pre-assembled polymerase complex to exist in the basal state. At these genes, the major regulatory checkpoint is therefore the licensing of the polymerase to enter a productive elongation phase, as opposed to the recruitment of polymerase itself. Signal-dependent elongation is a well-conserved regulatory mechanism, having been well-documented at the heat-shock genes, and more recently at rapidly induced immune genes in *Drosophila* (69, 70). The inhibition of BRD4 by highly selective chemical compounds, however, revealed that such an inhibition can selectively inhibit low-CpG containing genes (71). Although BRD4 activity is essential for a broad class of NF- $\kappa$ B dependent genes, the selective effect of the I-BET chemical antagonists implies multiple mechanisms of BRD4 recruitment. Indeed, there is evidence that acetylated NF- $\kappa$ B can directly recruit BRD4 (72).

#### CHROMATIN DYNAMICS AT ENHANCER REGIONS: A MECHANISM FOR IMMUNOLOGICAL MEMORY AND VARIABILITY?

Chromatin dynamics at promoters play a central role in how inducible genes are regulated, but advances in genome-wide studies have shed new light on how enhancer regions can also play a dynamic regulatory role. Because enhancer elements can be located great distances from the transcriptional start sites (ranging from a few kilobases to over a megabase), identification of bona fide enhancers has been a significant challenge, and the mechanisms of enhancer-mediated regulation of transcription remain somewhat unclear. As discussed above, enhancers are important in establishing the transcriptional competence of a locus following the binding of lineage-specifying factors. A highly multiplexed ChIP-seq study by Garber et al. in dendritic cells further developed this model, differentiating enhancer-binding transcription factors into three major regulatory classes: stably bound pioneer factors C/EBP and PU.1, which presumably bind to inaccessible chromatin early in lineage commitment; “primer” factors, such as Jun and AP-1, which are stably bound subsequent to pioneer factor binding and presumably contribute to local chromatin modification and remodeling; and finally the inducible factors, chiefly NF- $\kappa$ B and STAT proteins (73). Concurrent with inducible transcription factor binding, enhancer chromatin modifications are also dynamic. Active enhancer regions in a broad variety of cell types are distinguished by high levels of H3K4 monomethylation, and a number of ChIP-seq studies have utilized this specific mark to discover novel enhancers at a genome-wide level (74, 75).

The H3K27 acetyl mark has recently been demonstrated as another important enhancer-associated modification (76). Genome-wide analyses of this mark in macrophages has enabled functional categorization of enhancer elements into constitutive, poised, latent, and repressed classes based on the dynamics of the two histone modifications and PU.1 binding. Constitutively active enhancers are marked by PU.1 binding and high levels of both H3K4m1 and H3K27ac that remain stable upon stimulation. Poised enhancers only contain PU.1 and H3K4m1, and H3K27 is dynamically acetylated upon stimulation. Latent enhancers are



devoid of these marks, which are deposited upon stimulation. These enhancers govern the expression of late-phase genes, and are established in a stimulus-specific manner. Specifically, activation via TNF and IL1 $\beta$ , potent activators of the NF- $\kappa$ B signaling pathway, results in latent enhancers that are bound by the lineage definer PU.1 and NF- $\kappa$ B. By contrast, interferon stimulation results in the appearance of a latent enhancer repertoire enriched for PU.1 and STAT/IRF binding sites. In this manner, cells that have been exposed to a specific stimulus retain a short-term epigenetic memory of that stimulus, facilitating a more rapid and efficient transcriptional response upon subsequent stimulation (77). On the other hand, repressive chromatin modifications can be deposited at regulatory elements upon an initial stimulus, thereby attenuating responses to a secondary stimulus, and causing the cell to become tolerized against the stimulus. At the level of chromatin, it has been shown that promoters of tolerizable genes become hypoacetylated and adopt an inaccessible chromatin structure and are subsequently hyporesponsive (78).

Given that enhancers play such an important modulatory role in the inducible gene program, there has been considerable interest in how genetic variation of regulatory elements can influence gene expression. Heinz and colleagues recently showed that strain-specific genetic variations at enhancers affected PU.1 and CEB/P binding in macrophages. Loss or gain of these lineage-determining factors was concomitant with alterations of H3K4me1 and H3K27ac levels, as well as the expression levels of nearby genes. Importantly, the variability in lineage-factor binding sites resulted in a loss of NF- $\kappa$ B binding at enhancers, much more so than variation of  $\kappa$ B sites themselves (79). The strong correlation between loss of lineage-factor binding, loss of chromatin modification, and loss of NF- $\kappa$ B, support the role of NF- $\kappa$ B as a multifunctional switch that can only bind to sites marked and poised by lineage-defining factors.

## FUTURE DIRECTIONS AND ADVANCES

In the 25 years since its discovery, NF- $\kappa$ B has become one of the most heavily studied transcription factors. This is of course not without cause, as it is well-appreciated that a number of disease states, particularly inflammatory disease states, are at some level due to dysregulated NF- $\kappa$ B signaling. This review has attempted to summarize the broad range of chromatin-regulated mechanisms that govern the specificity of NF- $\kappa$ B-mediated transcription, the variety of ways NF- $\kappa$ B itself can influence chromatin structure, and a survey of how epigenetic programs initiated in response to NF- $\kappa$ B are established and propagated. A recurring theme in the study of NF- $\kappa$ B is that this transcription factor rarely acts alone. Lineage-determining transcription factors set the stage at the chromatin level, thereby dictating the response NF- $\kappa$ B will induce. Furthermore, a number of chromatin changes must take place at NF- $\kappa$ B dependent genes in order for proper gene induction to occur. In this sense, NF- $\kappa$ B can be thought of as the final regulatory switch. At rapidly induced genes, the switch controls late events such as the licensing of a pre-assembled polymerase into a productive elongation phase. In late-expressed genes, chromatin remodeling must occur, polymerase must be recruited and transition into an elongation phase takes place. This regulatory scheme largely focuses on events at the promoter, and indeed only fairly recently have the

mechanisms of enhancer regulation been appreciated. However, a number of outstanding questions remain to be answered. Post-translational modification of NF- $\kappa$ B has long been appreciated as an important regulatory mechanism, dictating cofactor specificity for active dimers, but the physiological necessity of many of these modifications remains unclear outside of *in vitro* settings. In many of these cases, the identity of modification-dependent interaction partners is unknown, and the spatiotemporal regulation of the majority of modifications remains unclear. Furthermore, the relative contributions of the multiple regulatory mechanisms remain unclear. For example, the p300/CBP bound by NF- $\kappa$ B plays a role in enhancing transcription of inducible genes (as in the case of the IFN $\beta$  enhanceosome), but its direct acetylation of NF- $\kappa$ B also plays a critical and perhaps more diverse regulatory role. It is possible that NF- $\kappa$ B cofactor requirements are dictated by cooperative transcription factor binding at specific regulatory elements. As with the case of the RelA Ser276 residue, generation of genetic models targeted at specific residues would be of tremendous value.

Genome-wide studies of the chromatin state during a pathogenic response have strengthened the understanding of the various regulatory mechanisms involved in establishing a competent transcriptional response. Although our understanding of the dynamics of NF- $\kappa$ B and its relationship with lineage-defining factors has deepened, it should be noted that there are numerous transcription factors and chromatin modifying complexes whose roles and relationship with NF- $\kappa$ B remain to be further characterized. Integrative and multiplexed studies have given us a snapshot of the different hierarchies of transcription factor binding (73), and similar studies examining the panoply of chromatin modifications will likely prove fruitful as well. By systematically targeting different chromatin modifiers, either chemically or genetically, a deeper understanding of the regulatory logic governing the NF- $\kappa$ B transcriptional response can be developed.

## ACKNOWLEDGMENTS

The work in the author's laboratory was supported by grants from the NIH (R37-AI33443 and RO1-AI066109).

## REFERENCES

- Hoffmann A, Baltimore D. Circuitry of nuclear factor kappaB signaling. *Immunol Rev* (2006) **210**:171–86. doi:10.1111/j.0105-2896.2006.00375.x
- Hayden MS, Ghosh S. Shared principles in NF-kappaB signaling. *Cell* (2008) **132**:344–62. doi:10.1016/j.cell.2008.01.020
- Carey M, Peterson CL, Smale ST. *Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (2008).
- Venters BJ, Pugh BE. How eukaryotic genes are transcribed. *Crit Rev Biochem Mol Biol* (2009) **44**:117–41. doi:10.1080/10409230902858785
- Malik S, Roeder RG. The metazoan mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* (2010) **11**:761–72. doi:10.1038/nrg2901
- Kawaji H, Frith MC, Katayama S, Sandelin A, Kai C, Kawai J, et al. Dynamic usage of transcription start sites within core promoters. *Genome Biol* (2006) **7**:R118. doi:10.1186/gb-2006-7-12-r118
- Juven-Gershon T, Hsu J-Y, Theisen JW, Kadonaga JT. The RNA polymerase II core promoter – the gateway to transcription. *Curr Opin Cell Biol* (2008) **20**:253–9. doi:10.1016/j.ccb.2008.03.003
- Smale ST, Kadonaga JT. The RNA polymerase II core promoter. *Annu Rev Biochem* (2003) **72**:449–79. doi:10.1146/annurev.biochem.72.121801.161520

9. Li B, Carey M, Workman JL. The role of chromatin during transcription. *Cell* (2007) **128**:707–19. doi:10.1016/j.cell.2007.01.015
10. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* (1997) **389**:251–60. doi:10.1038/38444
11. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res* (2011) **21**:381–95. doi:10.1038/cr.2011.22
12. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Res* (2011) **21**:564–78. doi:10.1038/cr.2011.42
13. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* (2010) **38**:576–89. doi:10.1016/j.molcel.2010.05.004
14. Ghisletti S, Barozzi I, Miettton F, Polletti S, De Santa F, Venturini E, et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* (2010) **32**:317–28. doi:10.1016/j.immuni.2010.02.008
15. Thanos D, Maniatis T. Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell* (1995) **83**:1091–100. doi:10.1016/0092-8674(95)90136-1
16. Agaloti T, Lomvardas S, Parekh B, Yie J, Maniatis T, Thanos D. Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell* (2000) **103**:667–78. doi:10.1016/S0092-8674(00)00169-0
17. Ford E, Thanos D. The transcriptional code of human IFN-beta gene expression. *Biochim Biophys Acta* (2010) **1799**:328–36. doi:10.1016/j.bbagr.2010.01.010
18. Plevy S, Gemberling JH, Hsu S, Donner AJ, Smale ST. Multiple control elements mediate activation of the murine and human interleukin 12 p40 promoters: evidence of functional synergy between C/EBP and Rel proteins. *Mol Cell Biol* (1997) **17**:4572–88.
19. Zhu C, Gagnidze K, Gemberling JH, Plevy S. Characterization of an activation protein-1-binding site in the murine interleukin-12 p40 promoter. Demonstration of novel functional elements by a reductionist approach. *J Biol Chem* (2001) **276**:18519–28. doi:10.1074/jbc.M100440200
20. Zhu C, Rao K, Xiong H, Gagnidze K, Li F, Horvath C, et al. Activation of the murine interleukin-12 p40 promoter by functional interactions between NFAT and ICSBP. *J Biol Chem* (2003) **278**:39372–82. doi:10.1074/jbc.M306441200
21. Murphy T, Cleveland M, Kulesza P, Magram J, Murphy K. Regulation of interleukin 12 p40 expression through an NF-kappa B half-site. *Mol Cell Biol* (1995) **15**:5258–67.
22. Weinmann AS, Plevy S, Smale ST. Rapid and selective remodeling of a positioned nucleosome during the induction of IL-12 p40 transcription. *Immunity* (1999) **11**:665–75. doi:10.1016/S1074-7613(00)80141-7
23. Weinmann AS, Mitchell DM, Sanjabi S, Bradley MN, Hoffmann A, Liou HC, et al. Nucleosome remodeling at the IL-12 p40 promoter is a TLR-dependent, Rel-independent event. *Nat Immunol* (2001) **2**:51–7. doi:10.1038/83168
24. Zhou L, Nazarian AA, Xu J, Tantin D, Corcoran LM, Smale ST. An inducible enhancer required for IL12b promoter activity in an insulated chromatin environment. *Mol Cell Biol* (2007) **27**:2698–712. doi:10.1128/MCB.00788-06
25. Saccani S, Pantano S, Natoli G. Two waves of nuclear factor kappaB recruitment to target promoters. *J Exp Med* (2001) **193**:1351–9. doi:10.1084/jem.193.12.1351
26. Müller CW, Rey FA, Sodeoka M, Verdine GL, Harrison SC. Structure of the NF-kappa B p50 homodimer bound to DNA. *Nature* (1995) **373**:311–7. doi:10.1038/373311a0
27. Chen FE, Huang D-B, Chen YQ, Ghosh G. Crystal structure of p50/p65 heterodimer of transcription factor NF-kappaB bound to DNA. *Nature* (1998) **391**:410–3. doi:10.1038/34356
28. Ghosh G, Wang VY-F, Huang D-B, Fusco A. NF- $\kappa$ B regulation: lessons from structures. *Immunol Rev* (2012) **246**:36–58. doi:10.1111/j.1600-065X.2012.01097.x
29. Steger DJ, Workman JL. Stable co-occupancy of transcription factors and histones at the HIV-1 enhancer. *EMBO J* (1997) **16**:2463–72. doi:10.1093/emboj/16.9.2463
30. Angelov D, Lenouvel F, Hans F, Müller CW, Bouvet P, Bednar J, et al. The histone octamer is invisible when NF-kappaB binds to the nucleosome. *J Biol Chem* (2004) **279**:42374–82. doi:10.1074/jbc.M407235200
31. Lone IN, Shukla MS, Charles Richard JL, Peshev ZY, Dimitrov S, Angelov D. Binding of NF- $\kappa$ B to nucleosomes: effect of translational positioning, nucleosome remodeling and linker histone H1. *PLoS Genet* (2013) **9**:e1003830. doi:10.1371/journal.pgen.1003830
32. Zhong H, May MJ, Jimi E, Ghosh S. The phosphorylation status of nuclear NF-kappa B determines its association with CBP/p300 or HDAC-1. *Mol Cell* (2002) **9**:625–36. doi:10.1016/S1097-2765(02)00477-X
33. Saccani S, Pantano S, Natoli G. Modulation of NF-kappaB activity by exchange of dimers. *Mol Cell* (2003) **11**:1563–74. doi:10.1016/S1097-2765(03)00227-2
34. Ramirez-Carrozzi VR, Nazarian AA, Li CC, Gore SL, Sridharan R, Imbalzano AN, et al. Selective and antagonistic functions of SWI/SNF and Mi-2 beta nucleosome remodeling complexes during an inflammatory response. *Genes Dev* (2006) **20**:282–96. doi:10.1101/gad.1383206
35. Ramirez-Carrozzi VR, Braas D, Bhatt DM, Cheng CS, Hong C, Doty KR, et al. A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* (2009) **138**:114–28. doi:10.1016/j.cell.2009.04.020
36. Hargreaves DC, Horng T, Medzhitov R. Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* (2009) **138**:129–45. doi:10.1016/j.cell.2009.05.047
37. Deaton AM, Bird AP. CpG islands and the regulation of transcription. *Genes Dev* (2011) **25**:1010–22. doi:10.1101/gad.2037511
38. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S, et al. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature* (2010) **464**:1082–6. doi:10.1038/nature08924
39. Lee J-H, Skalik DG. CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. *J Biol Chem* (2005) **280**:41725–31. doi:10.1074/jbc.M508312200
40. Fenouil R, Cauchy P, Koch F, Descostes N, Cabeza JZ, Innocenti C, et al. CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. *Genome Res* (2012) **22**:2399–408. doi:10.1101/gr.138776.112
41. O'Shea JM, Perkins ND. Regulation of the RelA (p65) transactivation domain. *Biochem Soc Trans* (2008) **36**:603–8. doi:10.1042/BST0360603
42. Vermeulen L, De Wilde G, Van Damme P, Vanden Berghe W, Haegeman G. Transcriptional activation of the NF-kappaB p65 subunit by mitogen- and stress-activated protein kinase-1 (MSK1). *EMBO J* (2003) **22**:1313–24. doi:10.1093/emboj/cdg139
43. Gerlo S, Kooijman R, Beck IM, Kolmus K, Spooren A, Haegeman G. Cyclic AMP: a selective modulator of NF- $\kappa$ B action. *Cell Mol Life Sci* (2011) **68**:3823–41. doi:10.1007/s00018-011-0757-8
44. Nihira K, Ando Y, Yamaguchi T, Kagami Y, Miki Y, Yoshida K. Pim-1 controls NF-kappaB signalling by stabilizing RelA/p65. *Cell Death Differ* (2010) **17**:689–98. doi:10.1038/cdd.2009.174
45. Zhong H, SuYang H, Erdjument-Bromage H, Tempst P, Ghosh S. The transcriptional activity of NF-kappaB is regulated by the IkappaB-associated PKAc subunit through a cyclic AMP-independent mechanism. *Cell* (1997) **89**:413–24. doi:10.1016/S0092-8674(00)80222-6
46. Zhong H, Voll RE, Ghosh S. Phosphorylation of NF-kappa B p65 by PKA stimulates transcriptional activity by promoting a novel bivalent interaction with the coactivator CBP/p300. *Mol Cell* (1998) **1**:661–71. doi:10.1016/S1097-2765(00)80066-0
47. Dong J, Jimi E, Zhong H, Hayden MS, Ghosh S. Repression of gene expression by unphosphorylated NF-kappaB p65 through epigenetic mechanisms. *Genes Dev* (2008) **22**:1159–73. doi:10.1101/gad.1657408
48. Mukherjee SP, Behar M, Birnbaum HA, Hoffmann A, Wright PE, Ghosh G. Analysis of the RelA:CBP/p300 interaction reveals its involvement in NF- $\kappa$ B-driven transcription. *PLoS Biol* (2013) **11**:e1001647. doi:10.1371/journal.pbio.1001647
49. Yang X-D, Huang B, Li M, Lamb A, Kelleher NL, Chen L-F. Negative regulation of NF-kappaB action by Set9-mediated lysine methylation of the RelA subunit. *EMBO J* (2009) **28**:1055–66. doi:10.1038/emboj.2009.55
50. Yang X-D, Tajkhorshid E, Chen L-F. Functional interplay between acetylation and methylation of the RelA subunit of NF-kappaB. *Mol Cell Biol* (2010) **30**:2170–80. doi:10.1128/MCB.01343-09
51. Kim J-W, Jang S-M, Kim C-H, An J-H, Kang E-J, Choi K-H. New molecular bridge between RelA/p65 and NF- $\kappa$ B target genes via histone acetyltransferase TIP60 cofactor. *J Biol Chem* (2012) **287**:7780–91. doi:10.1074/jbc.M111.278465
52. Huang B, Yang X-D, Lamb A, Chen L-F. Posttranslational modifications of NF-kappaB: another layer of regulation for NF-kappaB signaling pathway. *Cell Signal* (2010) **22**:1282–90. doi:10.1016/j.cellsig.2010.03.017

53. Perkins ND. Post-translational modifications regulating the activity and function of the nuclear factor kappa B pathway. *Oncogene* (2006) **25**:6717–30. doi:10.1038/sj.onc.1209937
54. van Essen D, Zhu Y, Saccani S. A feed-forward circuit controlling inducible NF- $\kappa$ B target gene activation by promoter histone demethylation. *Mol Cell* (2010) **39**:750–60. doi:10.1016/j.molcel.2010.08.010
55. Saccani S, Natoli G. Dynamic changes in histone H3 Lys 9 methylation occurring at tightly regulated inducible inflammatory genes. *Genes Dev* (2002) **16**:2219–24. doi:10.1101/gad.232502
56. De Santa F, Narang V, Yap ZH, Tusi BK, Burgold T, Austenaa L, et al. Jmjd3 contributes to the control of gene expression in LPS-activated macrophages. *EMBO J* (2009) **28**:3341–52. doi:10.1038/emboj.2009.271
57. Kruidenier L, Chung C, Cheng Z, Liddle J, Che K, Joberty G, et al. A selective jumoni H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* (2012) **488**:404–8. doi:10.1038/nature11262
58. Ghisletti S, Huang W, Jepsen K, Benner C, Hardiman G, Rosenfeld MG, et al. Cooperative NCoR/SMRT interactions establish a corepressor-based strategy for integration of inflammatory and anti-inflammatory signaling pathways. *Genes Dev* (2009) **23**:681–93. doi:10.1101/gad.1773109
59. Huang W, Ghisletti S, Perissi V, Rosenfeld MG, Glass CK. Transcriptional integration of TLR2 and TLR4 signaling at the NCoR derepression checkpoint. *Mol Cell* (2009) **35**:48–57. doi:10.1016/j.molcel.2009.05.023
60. Stender JDD, Pascual G, Liu W, Kaikkonen MU, Do K, Spann NJ, et al. Control of proinflammatory gene programs by regulated trimethylation and demethylation of histone H4K20. *Mol Cell* (2012) **48**:28–38. doi:10.1016/j.molcel.2012.07.020
61. Li P, Spann NJ, Kaikkonen MU, Lu M, Oh DY, Fox JN, et al. NCoR repression of LXRs restricts macrophage biosynthesis of insulin-sensitizing omega 3 fatty acids. *Cell* (2013) **155**:200–14. doi:10.1016/j.cell.2013.08.054
62. Chen X, Barozzi I, Termanini A, Prosperini E, Recchiuti A, Dalli J, et al. Requirement for the histone deacetylase Hdac3 for the inflammatory gene expression program in macrophages. *Proc Natl Acad Sci U S A* (2012) **109**:E2865–74. doi:10.1073/pnas.1121131109
63. Saunders A, Core LJ, Lis JT. Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* (2006) **7**:557–67. doi:10.1038/nrm1981
64. Fuda NJ, Ardehali MB, Lis JT. Defining mechanisms that regulate RNA polymerase II transcription in vivo. *Nature* (2009) **461**:186–92. doi:10.1038/nature08449
65. Brès V, Yoh SM, Jones KA. The multi-tasking P-TEFb complex. *Curr Opin Biol* (2008) **20**:334–40. doi:10.1016/j.ceb.2008.04.008
66. Nechaev S, Adelman K. Pol II waiting in the starting gates: regulating the transition from transcription initiation into productive elongation. *Biochim Biophys Acta* (2011) **1809**:34–45. doi:10.1016/j.bbagr.2010.11.001
67. Ainbinder E, Amir-Zilberstein L, Yamaguchi Y, Handa H, Dikstein R. Elongation inhibition by DRB sensitivity-inducing factor is regulated by the A20 promoter via a novel negative element and NF- $\kappa$ B. *Mol Cell Biol* (2004) **24**:2444–54. doi:10.1128/MCB.24.6.2444–2454.2004
68. Adelman K, Kennedy MA, Nechaev S, Gilchrist DA, Muse GW, Chinenov Y, et al. Immediate mediators of the inflammatory response are poised for gene activation through RNA polymerase II stalling. *Proc Natl Acad Sci U S A* (2009) **106**:18207–12. doi:10.1073/pnas.0910177106
69. Gilchrist DA, Fromm G, dos Santos G, Pham LN, McDaniel IE, Burkholder A, et al. Regulating the regulators: the pervasive effects of Pol II pausing on stimulus-responsive gene networks. *Genes Dev* (2012) **26**:933–44. doi:10.1101/gad.187781.112
70. Adelman K, Rogatsky I. RNA polymerase II stalling mediates cytokine gene expression. *Cell Cycle* (2010) **9**:630–1. doi:10.4161/cc.9.4.10841
71. Nicodeme E, Jeffrey KL, Schaefer U, Beinke S, Dewell S, Chung C, et al. Suppression of inflammation by a synthetic histone mimic. *Nature* (2010) **468**:1119–23. doi:10.1038/nature09589
72. Huang B, Yang X-D, Zhou M-M, Ozato K, Chen L-F. Brd4 coactivates transcriptional activation of NF- $\kappa$ B via specific binding to acetylated RelA. *Mol Cell Biol* (2009) **29**:1375–87. doi:10.1128/MCB.01365-08
73. Garber M, Yosef N, Goren A, Raychowdhury R, Thielke A, Guttman M, et al. A high-throughput chromatin immunoprecipitation approach reveals principles of dynamic gene regulation in mammals. *Mol Cell* (2012) **47**:810–22. doi:10.1016/j.molcel.2012.07.030
74. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* (2009) **459**:108–12. doi:10.1038/nature07829
75. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* (2009) **457**:854–8. doi:10.1038/nature07730
76. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* (2011) **470**:279–83. doi:10.1038/nature09692
77. Ostuni R, Piccolo V, Barozzi I, Polletti S, Termanini A, Bonifacio S, et al. Latent enhancers activated by stimulation in differentiated cells. *Cell* (2013) **152**:157–71. doi:10.1016/j.cell.2012.12.018
78. Foster SL, Hargreaves DC, Medzhitov R. Gene-specific control of inflammation by TLR-induced chromatin modifications. *Nature* (2007) **447**:972–8. doi:10.1038/nature05836
79. Heinz S, Romanoski CE, Benner C, Allison KA, Kaikkonen MU, Orozco LD, et al. Effect of natural genetic variation on enhancer selection and function. *Nature* (2013) **503**:487–92. doi:10.1038/nature12615

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 January 2014; paper pending published: 05 February 2014; accepted: 10 February 2014; published online: 25 February 2014.

Citation: Bhatt D and Ghosh S (2014) Regulation of the NF- $\kappa$ B-mediated transcription of inflammatory genes. *Front. Immunol.* 5:71. doi: 10.3389/fimmu.2014.00071  
This article was submitted to B Cell Biology, a section of the journal *Frontiers in Immunology*.

Copyright © 2014 Bhatt and Ghosh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# CIITA and its dual roles in MHC gene transcription

**Ballachanda N. Devaiah and Dinah S. Singer \***

Experimental Immunology Branch, National Cancer Institute, NIH, Bethesda, MD, USA

**Edited by:**

Ananda L. Roy, Tufts University School of Medicine, USA

**Reviewed by:**

John D. Colgan, The University of Iowa, USA

Barbara Nikolajczyk, Boston University School of Medicine, USA

**\*Correspondence:**

Dinah S. Singer, Experimental Immunology Branch, National Cancer Institute, NIH, 9000 Rockville Pike, Building 10, Room 4B36, Bethesda, MD 20892, USA  
e-mail: dinah.singer@nih.gov

Class II transactivator (CIITA) is a transcriptional coactivator that regulates  $\gamma$ -interferon-activated transcription of Major Histocompatibility Complex (MHC) class I and II genes. As such, it plays a critical role in immune responses: CIITA deficiency results in aberrant MHC gene expression and consequently in autoimmune diseases such as Type II bare lymphocyte syndrome. Although CIITA does not bind DNA directly, it regulates MHC transcription in two distinct ways – as a transcriptional activator and as a general transcription factor. As an activator, CIITA nucleates an enhanceosome consisting of the DNA binding transcription factors RFX, cyclic AMP response element binding protein, and NF- $\kappa$ B. As a general transcription factor, CIITA functionally replaces the TFIID component, TAF1. Like TAF1, CIITA possesses acetyltransferase (AT) and kinase activities, both of which contribute to proper transcription of MHC class I and II genes. The substrate specificity and regulation of the CIITA AT and kinase activities also parallel those of TAF1. In addition, CIITA is tightly regulated by its various regulatory domains that undergo phosphorylation and influence its targeted localization. Thus, a complex picture of the mechanisms regulating CIITA function is emerging suggesting that CIITA has dual roles in transcriptional regulation which are summarized in this review.

**Keywords: CIITA, MHC transcription, NLR/CATERPILLER proteins, enhanceosome, TAF1, general transcription factors**

The class II transactivator (CIITA) is a master regulator of major histocompatibility complex (MHC) gene expression. It induces *de novo* transcription of MHC class II genes and enhances constitutive MHC class I gene expression. The role of CIITA in regulating MHC gene transcription is well established: CIITA deficiency or aberrant expression is linked to the Type II bare lymphocyte syndrome and to cancer, respectively (1, 2). Although CIITA has been primarily characterized as a transcriptional regulator of MHC genes, it also regulates transcription of over 60 immunologically important genes, including interleukin 4 (IL-4), IL-10, and a variety of thyroid-specific genes (3, 4). Despite the important role of CIITA in regulating expression of these genes, the actual mechanisms by which it functions are still being unraveled. Early studies established that CIITA is a coactivator that nucleates the formation of an enhanceosome with transcription factors binding to enhancer elements in the upstream regions of the MHC genes (5, 6). In addition, more recent studies from our lab have demonstrated that it also functions as a component of the basal transcriptional machinery (7–9). Here we review the two distinct mechanisms by which CIITA regulates transcription of MHC class I and II genes and speculate on how these activities may be interconnected.

## CIITA BELONGS TO THE FAMILY OF NLR/CATERPILLER PROTEINS

The members of the NLR/CATERPILLER family of proteins are defined by their structures which include both a nucleotide binding domain (NBD) and a C-terminal leucine rich region (LRR) (10). Of the nearly two dozen genes within the family, many encode proteins that mediate inflammatory responses and whose aberrant

expression has been correlated with a variety of diseases (11). Of these, only CIITA is a critical mediator of adaptive immunity. Constitutive expression of CIITA is limited to antigen presenting cells. However,  $\gamma$ -interferon exposure induces *de novo* CIITA expression in most cell types (12, 13). Its activity is known to be modulated by several posttranslational modifications including phosphorylation, ubiquitination, acetylation, and deacetylation (14). CIITA can also self-associate and oligomerize (15, 16). CIITA activity is further modulated by its cellular localization: the GTP binding domain (GBD) of CIITA regulates its shuttling between the nucleus and cytoplasm (17, 18).

Class II transactivator consists of a series of regulatory domains that include an activation domain (AD), an acetyltransferase (AT) domain, a proline/serine/threonine (PST) domain, a GBD and finally, the canonical LRR domain common to all NLR proteins. The CIITA AD domain binds to general transcription factors and the CREB-binding protein (CBP), leading to activation of the MHC class II promoter and repression of the IL-4 promoter (19, 20). This domain also partially overlaps the region required for the AT activity of CIITA (7). The role of the PST domain, which while essential for CIITA function, remains unknown (21). The LRR domain, which interacts with the GBD, is known to play an important role in CIITA movement into the nucleus and in regulating its transactivation function (15, 21, 22). The GBD domain, which is perhaps the best studied among the CIITA domains, has been shown to be the site of interaction of several DNA binding transactivators (23). The GBD regulates translocation of CIITA: mutation or deletion of the GBD results in increased nuclear export, suggesting that it is a negative regulator of CIITA nuclear export (18). Two nuclear localization signals have been mapped to the N-terminal

domain, as well as an additional one in the C-terminus (18). CIITA also contains two LxxLL motifs which are crucial for the transactivation function and self-association of CIITA (24). In addition to these multifunctional domains, CIITA also contains two degrons in the AD and PST domains respectively. These degrons signal the rapid degradation of CIITA through the ubiquitin-proteasome pathway, and are responsible for the instability and short half-life of CIITA (~30 min) (25). CIITA function is also regulated through its oligomerization, which has been shown to be mediated by amino acids 253–410 (16). Each of the domains and regions listed above are also the targets of multiple post translational modifications which control their function (14). The expression of the CIITA gene itself is regulated by a complex mechanism involving four promoters and five enhancers that combine to form a dynamic chromatin structure (26, 27). Thus, CIITA is an extremely complex, unstable, and short lived protein, suggesting that its presence and function are tightly regulated.

### CIITA FUNCTIONS AS A TRANSACTIVATOR BY NUCLEATING AN ENHANCEOSOME

The role of CIITA as a central component of an enhanceosome has been characterized primarily for the MHC genes. MHC class II genes are transcriptionally controlled via conserved *cis*-acting elements in their promoters. These elements, named the W/S, X, X2, and Y boxes interact with specific *trans*-activating DNA binding factors to regulate MHC transcription either positively or negatively (28). The DNA binding factors involved, namely RFX (a hetero-multimer consisting of subunits RFX5, RFX-ANK, and RFX-AP), cyclic AMP response element binding protein (CREB), and NF-Y (A, B, and C), bind directly to the X, X2, and Y boxes respectively. These factors are constitutively expressed but their binding to the class II gene is not sufficient to support expression (29). CIITA interacts with each of these DNA binding factors (30). Those interactions depend on distinct structural domains within CIITA. CIITA transactivator function is dependent on the AD, PST, GBD, and LRR domains within its primary protein structure. RFX-ANK and NF-YC bind to the N-terminal acidic AD, whereas the remaining *trans*-activating factors interact with the GBD (23). The interaction of CIITA with the DNA-bound transcription factors serves to form a transcriptionally active complex, or enhanceosome (5). Importantly, the cognate DNA binding sites are spaced in a manner that supports the formation of an enhanceosome complex anchored by CIITA (31).

Within the enhanceosome, CIITA functions in the recruitment of various histone modifying enzymes, both activating and repressive. During MHC gene activation, CIITA recruits histone modifying enzymes such as the ATs p300, CBP, and the p300/CBP-associated factor (PCAF), as well as the methyltransferase CARM1 which function to support active transcription (32). In contrast, during MHC gene silencing, CIITA and RFX recruit and bind histone deacetylases HDAC1, HDAC2, and HDAC4 (31). CIITA also interacts with chromatin remodeling factors, such as BRG1 (33), and other coactivators such as SRC-1 (32). Thus, while CIITA does not directly bind DNA, it serves to nucleate and coordinate the various transcription factors and chromatin modifying enzymes that are necessary to support transcription of the class I and II genes.

The concept of a CIITA-nucleated enhanceosome fits well with the regulation of MHC class I and II transcription, and is supported by considerable circumstantial evidence. Such evidence includes the demonstration of direct interaction between CIITA and the other components of the predicted enhanceosome (5). However, CIITA also transactivates many genes that may not possess the *cis*-elements and *trans*-factors that are found on MHC genes (3). This raises the possibility that CIITA may also have a more direct function in transcription besides nucleating an enhanceosome. Indeed, as will be discussed below, CIITA functions as a component of the basal transcription machinery.

### CIITA FUNCTIONS AS A GENERAL TRANSCRIPTION FACTOR AND FUNCTIONAL HOMOLOG OF TAF1

The basal transcriptional machinery requires the assembly on the core promoter of a pre-initiation complex (PIC) that plays a central role in regulating transcription initiation. PIC assembly is initiated by the binding of the TFIID general transcription factor complex – composed of the TATA-binding protein (TBP) and a set of TBP-associated factors (TAFs) – to the core promoter (34). The largest factor in the TFIID complex, TAF1, has AT and kinase activities both of which are essential for transcription initiation (35, 36). In early studies from our lab, it was demonstrated that constitutive transcription of MHC class I genes depends on the AT activity of TAF1: MHC class I expression was abrogated at restrictive temperatures in TAF1 temperature-sensitive mutants. In contrast, CIITA-activated transcription of MHC genes was unaffected under these conditions, and thus independent of TAF1 function (37). This finding led to our discovery that CIITA, like TAF1, has an intrinsic AT activity and can bypass the requirement for TAF1 (7). Based on these findings, we proposed that CIITA functions as a general transcription factor that can substitute for TAF1 function during  $\gamma$ -interferon-activated MHC transcription (38).

Consistent with the model that CIITA is a general transcription factor that assembles a TFIID-like complex, CIITA is known to recruit and directly interact with components of the TFIID complex, including TBP, TAF6, and TAF9, as well as PTEFb and TFIIB which are components of the PIC (39, 40). Further supporting this model, we found that CIITA interacts with the TFIID component TAF7 (8) which acts as a check-point regulator of constitutive class I transcription initiation by inhibiting TAF1 AT activity (41). TAF7 binds directly to the region encompassing the AT domain of CIITA. Importantly, the binding of TAF7 to CIITA *in vitro* inhibits both its AT activity and transcription. *In vivo* knock-down of TAF7 resulted in a significant increase in CIITA-activated MHC class I gene expression (8). Taken together, these findings suggested that CIITA is a functional homolog of TAF1.

Further evidence for the parallels between CIITA and TAF1 came from the finding that CIITA, like TAF1, is a kinase (9). In the canonical TFIID complex, TAF1 is associated with and inhibited by TAF7 until PIC assembly is complete. Transcription initiation requires the release of TAF7 from TAF1, thereby revealing TAF1's essential AT activity. The release is mediated by TAF1 autophosphorylation by its intrinsic kinase activity, which is essential for initiating transcription (42). Interestingly, although TAF1 also phosphorylates TAF7, this does not cause the release of TAF7 from the TFIID complex. Rather, it modulates the subsequent



regulation of TFIID, BRD4, and PTEFb transcription factors by TAF7 (43, 44). The finding that CIITA AT activity bypasses the requirement for TAF1 during activated MHC class I transcription and that TAF7 inhibits this activity suggested that CIITA also might have a kinase activity responsible for dissociating TAF7. Indeed, we recently found that CIITA has intrinsic kinase activity (9). CIITA, like TAF1, autophosphorylates. This autophosphorylation prevents the binding of TAF7 which would otherwise inhibit its AT activity. CIITA phosphorylates TAF7, although at sites distinct from those phosphorylated by TAF1 (9). It remains to be determined whether TAF7 phosphorylation by CIITA, like by TAF1, modulates the subsequent regulation of TFIID and PTEFb by TAF7.

Similar to TAF1, the AT activity of CIITA was identified by its ability to acetylate histones *in vitro* (7). However, the actual *in vivo* substrate(s) for both CIITA and TAF1 remain to be defined. This leaves open the possibility that CIITA acetylates any of its numerous interacting partners *in vivo*. Additionally, the AT activity of CIITA is modulated by its own GBD; the binding of GTP at this site increases AT activity and nuclear localization of CIITA (7). Acetylation of CIITA on its N-terminal nuclear localization domain by PCAF is also known to be a signal for its nuclear localization (45). However, the loss of its AT domain does not affect CIITA's nuclear localization (18). Thus, while it is established that the AT activity of CIITA is undoubtedly required for its function in transcriptional regulation, many questions remain regarding the targets of this activity, the complexity of its regulation both directly by trans factors and indirectly by other regulatory domains on CIITA.

Class II transactivator function is also regulated by phosphorylation, as has been well documented (16, 46, 47). CIITA is known to be phosphorylated by PKA, PKC, GSK3, CK1, ERK1/2, and CKII kinases at various sites spanning its AT, PST, and LRR domains (14). These phosphorylation events are known to regulate its transactivation function, nuclear localization, oligomerization, and its ability to interact with DNA binding transactivators. The newly discovered kinase activity of CIITA adds an additional layer of complexity to the functional mechanisms by which CIITA operates. The kinase activity of CIITA and its ability to autophosphorylate are thus likely to have significant ramifications on its function. Indeed, we found that autophosphorylation of CIITA enhances its AT activity *in vitro*, suggesting that the two enzymatic activities of CIITA are interconnected. In contrast to the AT activity of CIITA, at least three independent substrates for CIITA kinase activity have been found thus far including TAF7, Histone H2B, and the TFIIF component RAP74 (9). While CIITA autophosphorylation regulates its ability to interact with TAF7 and consequently its AT activity, the purpose of TAF7 phosphorylation by CIITA is yet to be discovered. By analogy with TAF1 (44), we speculate that it modulates TAF7 binding to its downstream targets, BRD4, PTEFb, and TFIID. RAP74 phosphorylation by TAF1 is thought to help in coordinating the functions of different components of the pre-initiation complex (36). Whether CIITA phosphorylation of RAP74 serves the same purpose remains to be seen. Similarly, the phosphorylation of Histone H2B at Ser36 has been demonstrated to play a significant role in regulating transcription during cell cycle progression and stress response (48, 49). Therefore, the ability of CIITA to phosphorylate Histone H2B Ser36 (9) supports

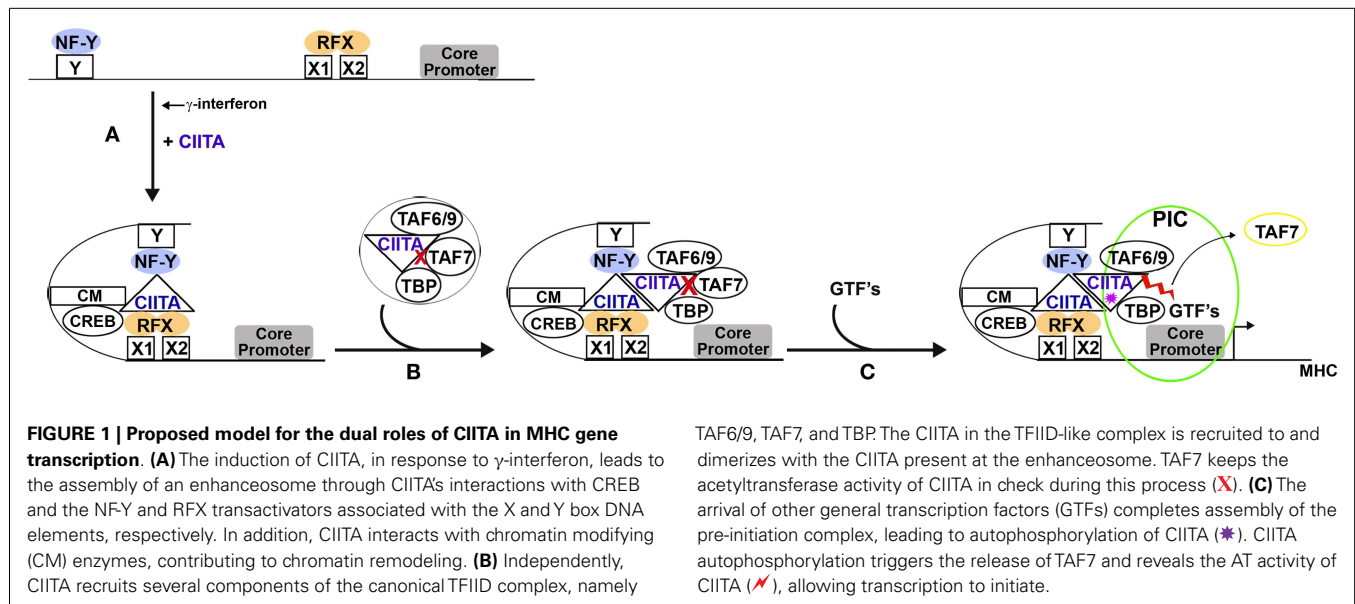
the idea that it has a role beyond its known function in regulating MHC genes. The substrate specificities of CIITA kinase activity, although very similar to those of TAF1, are distinct. CIITA phosphorylates TAF7 at a different site and phosphorylates all histones unlike TAF1 (9). We speculate that the distinct nature of CIITA kinase activity suggests that it may have other substrates, possibly in the enhanceosome complex that it nucleates.

## INTEGRATING THE DUAL FUNCTIONS OF CIITA: A MODEL

Although CIITA has no obvious structural homology with TAF1, it has remarkable functional parallels. Like TAF1, CIITA associates with general transcription factors to form a TFIID-like complex (8, 39, 40). Both CIITA and TAF1 have AT activity that is essential for transcription (7, 35, 41); both AT activities are regulated by TAF7 (8, 41). CIITA and TAF1 both have kinase activity that results in autophosphorylation leading to TAF7 dissociation (9, 36, 42). Taken together, these findings demonstrate that CIITA functions as a general transcription factor and functional homolog of TAF1, independent of its role as a coactivator that nucleates an enhanceosome. The dual functionality of CIITA, as a coactivator and a general transcription factor, leads to the question of how these two disparate functions are coordinated during MHC gene regulation. We propose the following model (Figure 1).

Following induction by  $\gamma$ -interferon, CIITA nucleates the formation of an enhanceosome through its interaction with transactivators that are constitutively bound to conserved DNA elements within the extended promoters of MHC class I and II (Figure 1A). We speculate that within the enhanceosome, CIITA may participate in chromatin remodeling, along with the histone modifying enzymes it recruits to the enhanceosome. Indeed, CIITA efficiently acetylates histones H3 and H4 (7). The acetylations may serve to maintain the MHC class I- and II-associated chromatin in a transcriptionally active conformation. Alternatively, within the enhanceosome, CIITA may acetylate any one of its interacting partners: the binding site on CIITA for many of the CIITA interacting partners in the MHC II enhanceosome (e.g., RFX-ANK and NF- $\kappa$ B) maps to the N-terminal  $\alpha$ -helical acidic domain between amino acids 58 and 94 (23), which is immediately adjacent to the CIITA AT domain between amino acids 94 and 132 (7). Whether these factors are substrates of the AT activity or whether their binding to CIITA within the context of the enhanceosome activates or represses AT activity remain intriguing questions. CIITA also efficiently phosphorylates all four histones (9). Phosphorylation of H2B by CIITA has been mapped to Ser36 (9), an event that leads to increased transcription and survival during cell stress (49). Thus it is possible that CIITA's enzymatic activities contribute to its coactivator function.

We further speculate that the TFIID-like complex assembled by CIITA is recruited to the promoters of MHC class I and II genes by the CIITA-nucleated enhanceosome through dimerization of the CIITA molecules contained within each complex (Figure 1B). According to this model, following nucleation by CIITA, the enhanceosome would interact with the TFIID-like CIITA complex and deliver it to the downstream core promoter. The binding of the TFIID-like CIITA complex to the core promoter would trigger the assembly of a PIC. Once PIC assembly is complete, CIITA would autophosphorylate, releasing TAF7 from the TFIID-like complex.



The release of TAF7 would relieve the inhibition of the AT activity associated with the CIITA in the TFIID-like complex. This AT activity, like that of TAF1, is required for transcription initiation.

As noted above, CIITA also activates transcription of non-MHC genes that do not support formation of enhanceosomes comprised of NF-Y and RFX (3). We speculate that these promoters recruit CIITA to the upstream regulatory regions of promoters through its binding to a variety of other transcription factors. This CIITA, by dimerization with CIITA in the TFIID-like complex is then able to deliver it to the core promoter. Future experiments will test this model.

The fact that CIITA is one of a family of NLR/CATERPILLER proteins also raises the possibility that other members of the family may have similar bi-functionality. Indeed, NLRC5, another NLR family member, regulates MHC class I gene transcription (50). NLRC5 domain structure is similar to that of CIITA; it has been proposed to assemble an enhanceosome on MHC class I promoters. It will be of interest to determine whether NLRC5 has enzymatic activities and dual roles in transcription analogous to those of CIITA.

## CONCLUSION

Class II transactivator regulates transcription of both MHC and non-MHC genes through its multiplicity of functions. It functions as a transcriptional transactivator in assembling an enhanceosome with transcription factors bound to the distal promoters of MHC genes and as a general transcription factor and functional homolog of TAF1 associated with the core promoter. While it is possible that these CIITA functions are independent, the cross-regulation of the AT and kinase activities of CIITA suggests that they are interconnected. Thus, we speculate that CIITA initially assembles an enhanceosome with DNA-bound transactivators, first recruiting histone modifying enzymes and then a TFIID-like complex containing CIITA that nucleates assembly of the PIC. Within the PIC, CIITA functionally replaces TAF1 to initiate transcription.

As other substrates for the enzymatic activities of CIITA are discovered in the future, it is likely that the complexity of CIITA function will further increase. Which genes CIITA activates as a basal transcription factor, and if this is determined by the presence or absence of upstream nucleation sites for an enhanceosome, will also be subjects for future research.

## ACKNOWLEDGMENTS

We thank Dr. Paul Roche for his critical review of this manuscript. We also thank members of our lab for helpful discussions and apologize to researchers whose work may not be cited due to lack of space. This research was supported by the Intramural Research Program of the Center for Cancer Research (CCR), National Cancer Institute, NIH.

## REFERENCES

- Steimle V, Otten LA, Zufferey M, Mach B. Complementation cloning of an MHC class II transactivator mutated in hereditary MHC class II deficiency (or bare lymphocyte syndrome). *Cell* (1993) 75(1):135–46. doi:10.1016/0092-8674(93)90685-J
- Vora AR, Rodgers S, Parker AJ, Start R, Rees RC, Murray AK. An immunohistochemical study of altered immunomodulatory molecule expression in head and neck squamous cell carcinoma. *Br J Cancer* (1997) 76(7):836–44. doi:10.1038/bjc.1997.472
- Nagarajan UM, Bushey A, Boss JM. Modulation of gene expression by the MHC class II transactivator. *J Immunol* (2002) 169(9):5078–88.
- Mori-Aoki A, Pietrarelli M, Nakazato M, Caturegli P, Kohn LD, Suzuki K. Class II transactivator suppresses transcription of thyroid-specific genes. *Biochem Biophys Res Commun* (2000) 278(1):58–62. doi:10.1006/bbrc.2000.3769
- Masternak K, Muhlethaler-Mottet A, Villard J, Zufferey M, Steimle V, Reith W. CIITA is a transcriptional coactivator that is recruited to MHC class II promoters by multiple synergistic interactions with an enhanceosome complex. *Genes Dev* (2000) 14(9):1156–66. doi:10.1101/gad.14.9.1156
- Fontes JD, Kanazawa S, Nekrep N, Peterlin BM. The class II transactivator CIITA is a transcriptional integrator. *Microbes Infect* (1999) 1(11):863–9. doi:10.1016/S1286-4579(99)00232-4
- Raval A, Howcroft TK, Weissman JD, Kirshner S, Zhu XS, Yokoyama K, et al. Transcriptional coactivator, CIITA, is an acetyltransferase that bypasses a promoter requirement for TAF(II)250. *Mol Cell* (2001) 7(1):105–15. doi:10.1016/S1097-2765(01)00159-9

8. Devaiah BN, Lu H, Geggion A, Sercan Z, Zhang H, Clifford RJ, et al. Novel functions for TAF7, a regulator of TAF1-independent transcription. *J Biol Chem* (2010) **285**(50):38772–80. doi:10.1074/jbc.M110.173864
9. Soe KC, Devaiah BN, Singer DS. Transcriptional coactivator CIITA, a functional homolog of TAF1, has kinase activity. *Biochim Biophys Acta* (2013) **1829**(11):1184–90. doi:10.1016/j.bbaggm.2013.09.001
10. Harton JA, Linhoff MW, Zhang J, Ting JP. Cutting edge: CATERPILLER: a large family of mammalian genes containing CARD, pyrin, nucleotide-binding, and leucine-rich repeat domains. *J Immunol* (2002) **169**(8):4088–93.
11. Lich JD, Ting JP. CATERPILLER (NLR) family members as positive and negative regulators of inflammatory responses. *Proc Am Thorac Soc* (2007) **4**(3):263–6. doi:10.1513/pats.200701-022AW
12. Chin KC, Mao C, Skinner C, Riley JL, Wright KL, Moreno CS, et al. Molecular analysis of G1B and G3A IFN gamma mutants reveals that defects in CIITA or RFX result in defective class II MHC and Ii gene induction. *Immunity* (1994) **1**(8):687–97. doi:10.1016/1074-7613(94)90039-6
13. Steimle V, Siegrist CA, Mottet A, Lisowska-Grospierre B, Mach B. Regulation of MHC class II expression by interferon-gamma mediated by the transactivator gene CIITA. *Science* (1994) **265**(5168):106–9. doi:10.1126/science.8016643
14. Wu X, Kong X, Luchsinger L, Smith BD, Xu Y. Regulating the activity of class II transactivator by posttranslational modifications: exploring the possibilities. *Mol Cell Biol* (2009) **29**(21):5639–44. doi:10.1128/MCB.00661-09
15. Linhoff MW, Harton JA, Cressman DE, Martin BK, Ting JP. Two distinct domains within CIITA mediate self-association: involvement of the GTP-binding and leucine-rich repeat domains. *Mol Cell Biol* (2001) **21**(9):3001–11. doi:10.1128/MCB.21.9.3001-3011.2001
16. Tosi G, Jabrane-Ferrat N, Peterlin BM. Phosphorylation of CIITA directs its oligomerization, accumulation and increased activity on MHCII promoters. *EMBO J* (2002) **21**(20):5467–76. doi:10.1093/emboj/cdf557
17. Harton JA, Cressman DE, Chin KC, Der CJ, Ting JP. GTP binding by class II transactivator: role in nuclear import. *Science* (1999) **285**(5432):1402–5. doi:10.1126/science.285.5432.1402
18. Raval A, Weissman JD, Howcroft TK, Singer DS. The GTP-binding domain of class II transactivator regulates its nuclear export. *J Immunol* (2003) **170**(2):922–30.
19. Fontes JD, Jiang B, Peterlin BM. The class II trans-activator CIITA interacts with the TBP-associated factor TAFII32. *Nucleic Acids Res* (1997) **25**(12):2522–8. doi:10.1093/nar/25.12.2522
20. Sisk TJ, Gourley T, Roys S, Chang CH. MHC class II transactivator inhibits IL-4 gene transcription by competing with NF-AT to bind the coactivator CREB binding protein (CBP)/p300. *J Immunol* (2000) **165**(5):2511–7.
21. Chin KC, Li GG, Ting JP. Importance of acidic, proline/serine/threonine-rich, and GTP-binding regions in the major histocompatibility complex class II transactivator: generation of transdominant-negative mutants. *Proc Natl Acad Sci U S A* (1997) **94**(6):2501–6. doi:10.1073/pnas.94.6.2501
22. Hake SB, Masternak K, Kammerbauer C, Janzen C, Reith W, Steimle V. CIITA leucine-rich repeats control nuclear localization, in vivo recruitment to the major histocompatibility complex (MHC) class II enhancosome, and MHC class II gene transactivation. *Mol Cell Biol* (2000) **20**(20):7716–25. doi:10.1128/MCB.20.20.7716-7725.2000
23. Harton JA, Ting JP. Class II transactivator: mastering the art of major histocompatibility complex expression. *Mol Cell Biol* (2000) **20**(17):6185–94. doi:10.1128/MCB.20.17.6185-6194.2000
24. Sisk TJ, Roys S, Chang CH. Self-association of CIITA and its transactivation potential. *Mol Cell Biol* (2001) **21**(15):4919–28. doi:10.1128/MCB.21.15.4919-4928.2001
25. Schnappauf F, Hake SB, Camacho Carvajal MM, Bontron S, Lisowska-Grospierre B, Steimle V. N-terminal destruction signals lead to rapid degradation of the major histocompatibility complex class II transactivator CIITA. *Eur J Immunol* (2003) **33**(8):2337–47. doi:10.1002/eji.200323490
26. Muhlethaler-Mottet A, Otten LA, Steimle V, Mach B. Expression of MHC class II molecules in different cellular and functional compartments is controlled by differential usage of multiple promoters of the transactivator CIITA. *EMBO J* (1997) **16**(10):2851–60. doi:10.1093/emboj/16.10.2851
27. Ni Z, Abou El Hassan M, Xu Z, Yu T, Bremner R. The chromatin-remodeling enzyme BRG1 coordinates CIITA induction through many interdependent distal enhancers. *Nat Immunol* (2008) **9**(7):785–93. doi:10.1038/ni.1619
28. Ting JP, Trowsdale J. Genetic control of MHC class II expression. *Cell* (2002) **109**(Suppl):S21–33. doi:10.1016/S0092-8674(02)00696-7
29. Choi NM, Majumder P, Boss JM. Regulation of major histocompatibility complex class II genes. *Curr Opin Immunol* (2011) **23**(1):81–7. doi:10.1016/j.coi.2010.09.007
30. Zhu XS, Linhoff MW, Li G, Chin KC, Maity SN, Ting JP. Transcriptional scaffold: CIITA interacts with NF-Y, RFX, and CREB to cause stereospecific regulation of the class II major histocompatibility complex promoter. *Mol Cell Biol* (2000) **20**(16):6051–61. doi:10.1128/MCB.20.16.6051-6061.2000
31. Wright KL, Ting JP. Epigenetic regulation of MHC-II and CIITA genes. *Trends Immunol* (2006) **27**(9):405–12. doi:10.1016/j.it.2006.07.007
32. Zika E, Ting JP. Epigenetic control of MHC-II: interplay between CIITA and histone-modifying enzymes. *Curr Opin Immunol* (2005) **17**(1):58–64. doi:10.1016/j.coi.2004.11.008
33. Mudhasani R, Fontes JD. Multiple interactions between BRG1 and MHC class II promoter binding proteins. *Mol Immunol* (2005) **42**(6):673–82. doi:10.1016/j.molimm.2004.09.021
34. Roeder RG. The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem Sci* (1996) **21**(9):327–35. doi:10.1016/0968-0004(96)10050-5
35. Mizzen CA, Yang XJ, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, et al. The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* (1996) **87**(7):1261–70. doi:10.1016/S0092-8674(00)81821-8
36. Dikstein R, Ruppert S, Tjian R. TAFII250 is a bipartite protein kinase that phosphorylates the base transcription factor RAP74. *Cell* (1996) **84**(5):781–90. doi:10.1016/S0092-8674(00)81055-7
37. Weissman JD, Howcroft TK, Singer DS. TAF(II)250-independent transcription can be conferred on a TAF(II)250-dependent basal promoter by upstream activators. *J Biol Chem* (2000) **275**(14):10160–7. doi:10.1074/jbc.275.14.10160
38. Howcroft TK, Raval A, Weissman JD, Geggion A, Singer DS. Distinct transcriptional pathways regulate basal and activated major histocompatibility complex class I expression. *Mol Cell Biol* (2003) **23**(10):3377–91. doi:10.1128/MCB.23.10.3377-3391.2003
39. Mahanta SK, Scholl T, Yang FC, Strominger JL. Transactivation by CIITA, the type II bare lymphocyte syndrome-associated factor, requires participation of multiple regions of the TATA box binding protein. *Proc Natl Acad Sci U S A* (1997) **94**(12):6324–9. doi:10.1073/pnas.94.12.6324
40. Kanazawa S, Okamoto T, Peterlin BM. Tat competes with CIITA for the binding to P-TEFb and blocks the expression of MHC class II genes in HIV infection. *Immunity* (2000) **12**(1):61–70. doi:10.1016/S1074-7613(00)80159-4
41. Geggion A, Weissman JD, Singer DS. TAFII55 binding to TAFII250 inhibits its acetyltransferase activity. *Proc Natl Acad Sci U S A* (2001) **98**(22):12432–7. doi:10.1073/pnas.211444798
42. Geggion A, Weissman JD, Zhou M, Brady JN, Singer DS. TAF7: a possible transcription initiation check-point regulator. *Proc Natl Acad Sci U S A* (2006) **103**(3):602–7. doi:10.1073/pnas.0510031103
43. Geggion A, Weissman JD, Lu H, Zhou M, Dasgupta A, Ribble R, et al. TFIID component TAF7 functionally interacts with both TFIIF and P-TEFb. *Proc Natl Acad Sci U S A* (2008) **105**(14):5367–72. doi:10.1073/pnas.0801637105
44. Devaiah BN, Singer DS. Cross-talk among RNA polymerase II kinases modulates C-terminal domain phosphorylation. *J Biol Chem* (2012) **287**(46):38755–66. doi:10.1074/jbc.M112.412015
45. Spilianakis C, Papamatheakis J, Kretsovali A. Acetylation by PCAF enhances CIITA nuclear accumulation and transactivation of major histocompatibility complex class II genes. *Mol Cell Biol* (2000) **20**(22):8489–98. doi:10.1128/MCB.20.22.8489-8498.2000
46. Sisk TJ, Nickerson K, Kwok RP, Chang CH. Phosphorylation of class II transactivator regulates its interaction ability and transactivation function. *Int Immunol* (2003) **15**(10):1195–205. doi:10.1093/intimm/dxg116
47. Li G, Harton JA, Zhu X, Ting JP. Downregulation of CIITA function by protein kinase A (PKA)-mediated phosphorylation: mechanism of prostaglandin E, cyclic AMP, and PKA inhibition of class II major histocompatibility complex expression in monocytic lines. *Mol Cell Biol* (2001) **21**(14):4626–35. doi:10.1128/MCB.21.14.4626-4635.2001
48. Maile T, Kwoczyński S, Katzenberger RJ, Wassarman DA, Sauer F. TAF1 activates transcription by phosphorylation of serine 33 in histone H2B. *Science* (2004) **304**(5673):1010–4. doi:10.1126/science.1095001

49. Bungard D, Fuerth BJ, Zeng PY, Faubert B, Maas NL, Viollet B, et al. Signaling kinase AMPK activates stress-promoted transcription via histone H2B phosphorylation. *Science* (2010) **329**(5996):1201–5. doi:10.1126/science.1191241
50. Meissner TB, Li A, Kobayashi KS. NLRC5: a newly discovered MHC class I trans-activator (CITA). *Microbes Infect* (2012) **14**(6):477–84. doi:10.1016/j.micinf.2011.12.007

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 24 October 2013; accepted: 07 December 2013; published online: 20 December 2013.

Citation: Devaiah BN and Singer DS (2013) CIITA and its dual roles in MHC gene transcription. *Front. Immunol.* **4**:476. doi: 10.3389/fimmu.2013.00476

This article was submitted to *B Cell Biology*, a section of the journal *Frontiers in Immunology*.

Copyright © 2013 Devaiah and Singer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.