Title
Mechanistically Understanding the Selective Regulation of Inflammatory Genes by NF-kB

Author
Daly, Allison

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Mechanistically Understanding the Selective Regulation of Inflammatory Genes by NF-κB

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular Biology

by

Allison Erysian Daly

2023

ABSTRACT OF THE DISSERTATION

Mechanistically Understanding the Selective
Regulation of Inflammatory Genes by NF-κB

by

Allison Erysian Daly
Doctor of Philosophy in Molecular Biology
University of California, Los Angeles, 2023
Professor Stephen T. Smale, Chair

Through billions of years of evolution, immune cells have developed diverse mechanisms to resolve pathogenic infection and heal injury. Proper immune cell activation includes receptor-ligand engagement, intracellular signaling cascades, and transcriptome-wide changes. Dysregulation of proinflammatory cytokines during infection can lead to severe complications. The NF-κB family regulates many pro-inflammatory cytokines associated with inflammation and immune system imbalance, such as *Tnfa, Il1b,* and *Il6*. Although NF-κB is one of the best-studied transcription factors, the molecular mechanisms underlying dimer-specific function remain largely unknown. Our research explores the role of dimers that are not essential for cellular or organismal survival but have distinct roles in immune cell activation. This research employs a macrophage model system to define dimer-specific roles of NF-κB and the molecular mechanisms underlying their regulation. Using a reductionist system enables us to explore dimer-specific functions with unprecedented depth in the cell context. We focused on c-Rel and p50 due to their potent activation downstream of viral and bacterial PRRs and their highly specific roles in immune cell

activation. We employ both genomic and biochemical approaches such as RNA-seq, ChIP-seq, Sequential ChIP-seq, EMSAs, and co-immunoprecipitations to uncover the mechanistic role of c-Rel and p50. This research provides novel insights into the regulatory logic employed by NF-κB by revealing the underlying mechanisms that make dimer-specific functions possible.

The dissertation of Allison Erysian Daly is approved.

Alexander Hoffmann

Man Hing Li

Michael F. Carey

Steven J. Bensinger

Stephen T. Smale, Committee Chair

University of California, Los Angeles

2023

DEDICATION

To my parents,

Barbara Erysian and Daniel Daly.

For always providing

encouragement,

guidance,

and inspiration.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

FIGURES

TABLES

CHAPTER 1

ACKNOWLEDGMENTS

I want to share my deep gratitude for the support, guidance, and encouragement I received from my advisor, Dr. Stephen Smale, the Smale lab, my friends, and my family during my Ph.D. Dr. Smale is a remarkable mentor who values the pursuit of rigorous science, thorough investigation, and the discovery of new scientific truths. He constantly inspired me to approach my research with creativity, open-mindedness, and meticulous attention to detail. His humility, passion for science, and thoughtful guidance continually inspired me to strive for excellence. Dr. Smale's dedication to being an amazing mentor was beyond what I could've imagined from a Ph.D. advisor. Thank you, Dr. Smale, for your selflessness, generosity, and amazing ability to explain complex concepts so I could understand them.

I am also grateful to have been able to learn and grow with a wonderful group of colleagues and scientists: Prabhat Prubey, Phil Scumpia, An-Chieh Feng, Jerry Hung-Hoa Lo, Fionna Shue, Amber Ruccia, Vasileios Ragkousis, Ariel Floro, Lijing Chen, Sofia Soltero, Camila Cordero, Lily Shang, and Joowon Seo. I have so many fond memories of us learning, struggling, and laughing together. You all brightened my days, taught me so much about science and life, and helped make the lab feel like my home away from home. I am also thankful for my committee, Alexander Hoffmann, Steven Bensinger, Melody Li, and Michael Carey who have helped guide my research, collaborated scientifically with me, and encouraged me throughout my Ph.D.

Finally, to my friends and family, thank you for your unending encouragement, support, and love. You helped give me perspective in moments when I was feeling overwhelmed by the challenges of the research; encouraged me to pursue my dreams; and celebrated all the small victories with me. Through conversation, meals, dancing, and laughing you brought me so much joy. I am incredibly thankful for my community in Los Angeles, my lifelong friends who live all over the world, and my dearest family.

Chapter 1 contains reprints of figures and tables from published manuscripts with granted permission from the publishers. Figure 1-1, Figure 1-2, Figure 1-3, and Table 1-2 are reprints granted by Springer Nature (License Numbers: 5521620667429, 5525540839250, 5521620922823, 5521621352912). Figure 1-4 and Table 1-1 are reprints granted by Elsevier (License Numbers: 5521621048157, 5521621217081).

Appendix A is a reprint of an article published in Nature Immunology.

Zhou, Q.D., Chi, X., Lee, M.S., Hsieh, W.Y., Mkrtchyan, J.J., Feng, A.C., He, C., York, A.G., Bui, V.L., Kronenberger, E.B., Ferrari, A., Xiao, X., **Daly, A.E**., Tarling, E.J., Damoiseaux, R., Scumpia, P.O., Smale, S.T., Williams, K.J., Tontonoz, P., Bensinger, S.J. (2020). Interferon-mediated reprogramming of membrane cholesterol to evade bacterial toxins. *Nat Immunol* ***21***, 746-755 (2020). DOI: 10.1038/s41590-020-0695-4

S.J.B. conceived the study. Q.D.Z. led the design and execution of experiments. X.C. codesigned and performed all flow cytometry experiments and data analysis. V.L.B., W.Y.H., and J.J.M. contributed to flow cytometry experiments. J.J.M. and M.S.L. contributed to protein purification and staining experiments. Q.D.Z., M.S.L., and R.D. developed and performed live-cell imaging assays. C.H. performed NanoSIMS analysis and contributed to protein purification. J.J.M. and E.B.K. contributed to RNA analysis. W.Y.H. and A.G.Y. performed GC–MS analysis with help from Q.D.Z. and E.B.K. K.J.W. conducted lipidomic studies. X.X., A.F., P.T., and E.J.T. contributed to *Abcg1*, *Abca1,* and SCAP KO studies. **A.E.D.** analyzed gene expression data. A.-C.F., P.O.S., M.S.L., and S.T.S. conceptualized and developed the in vivo SLO challenge assay. W.Y.H. contributed to data visualization. S.J.B., Q.D.Z., W.Y.H., X.C., and P.T. contributed to the construction of the manuscript.

Appendix B is in preparation for publication.

Feng A.C., Thomas B., Liu X., **Daly A.E.,** Sun F., Chen L.J., Carey M.F., Scumpia P.O. and Smale S.T. NF-κB Broadly Orchestrates Nucleosome Remodeling During the Primary Response to Toll-Like Receptor 4 Signaling.

S.T.S. oversaw this study and wrote the manuscript. A.C.F., B.J.T., F.S., L.C., and X.L. designed and performed experiments. **A.E.D** generated the ATAC-seq treated with MAPK inhibitor. M.F.C., P.O.S., and S.T.S. provided guidance and input into the experimental design and data interpretation.

All the work described in the dissertation was directed by Stephen Smale.

The work in this thesis was supported by the UCLA T32 Microbial Pathogenesis Training Grant, the Whitcome Pre-Doctoral Fellowship in Molecular Biology, and the NIH F31 Ruth L. Kirschstein Predoctoral National Research Service Award from the National Institute of Allergy and Infectious Disease.

VITA

2017	Bachelor of Arts, Molecular and Cell Biology
	Minor, Chemistry
	University of California, Berkeley
2018 – 2019	Teaching Assistant
	Introduction to Microbiology
	Department of Microbiology, Immunology, and Molecular Genetics
	University of California, Los Angeles
2019 – 2021	Molecular Pathogenesis Training Grant Recipient
	University of California, Los Angeles
2021 – 2022	Whitcome Fellowship Recipient
	University of California, Los Angeles
2021 – 2023	NIH F31 Ruth L. Kirschstein Predoctoral National Research Service Award
	National Institute of Allergy and Infectious Disease

Publications

Zhou Q.D., Chi X., Lee M.S., Hsieh W.Y., Mkrtchyan J.J., Feng A.C., He C., York A.G., Bui V.L., Kronenberger E.B., Ferrari A, Xiao X., **Daly A.E.,** Tarling E.J., Damoiseaux R., Scumpia P.O., Smale S.T., Williams K.J., Tontonoz P. and Bensinger S.J. Interferon-mediated reprogramming of membrane cholesterol to evade bacterial toxins. *Nat Immunology.* (2020) doi:10.1038/s41590-020-0695-4

Presentations

September 2022	Mechanistically understanding the selective regulation of inflammatory genes by NF-κB. Poster Presentation, Cell Symposia: The Neuro-Immune Axis. Lisbon, Portugal.
August 2022	Mechanistically understanding the selective regulation of inflammatory genes by NF-κB. Poster Presentation, EMBL: Transcription and Chromatin. Heidelberg, Germany.
March 2022	Mechanistically understanding the NF-κB immune response to bacterial infection. Poster Presentation, 61st Midwinter Conference for Immunologists. Asilomar, California.
March 2021	Understanding selective NF-κB responses to bacterial infection. Research in Progress Talk, UCLA Molecular Pathogenesis Training Grant. Los Angeles, California.
February 2020	Mechanistically understanding NF-κB regulation of inflammatory genes. Research in Progress Talk, UCLA David Geffen School of Medicine. Los Angeles, California.
January 2020	The selective regulation of inflammatory genes by NF-κB. Poster Presentation, UCLA Molecular Biology Institute Retreat. Los Angeles, California.
March 2017	Investigating the role of transcription factors in anti-inflammatory M2 macrophage polarization, adipose tissue homeostasis, and type 2 diabetes. Poster Presentation, UC Berkeley Honors Symposia. Berkeley, California.

CHAPTER 1

Introduction:

The Activation of Immune Cells Through Transcriptional Regulation

THE IMMUNE SYSTEM

The immune system has a remarkable ability to detect changes in the surrounding environment and initiate an appropriate response to pathogenic stimuli. A major challenge for the immune system is resolving infection and quickly returning to homeostasis. Through billions of years of evolution, the immune system has developed diverse mechanisms to quickly recognize and respond to foreign invaders[1]. These mechanisms allow the immune system to effectively distinguish self from non-self and clear pathogenic infection. Since the natural environment contains an abundance of bacterial, viral, and other pathogenic species, multicellular organisms depend on the immune system for survival.

The immune system is divided into two categories, innate and adaptive immune responses. In mammals, the innate and adaptive immune systems work together to clear infection. All immune cells develop from hematopoietic stem cells (HSCs) which reside in the bone marrow. Through organized processes, HSCs give rise to lymphoid progenitor and myeloid progenitor cells[2]. Myeloid progenitor cells terminally differentiate into innate immune cells, which include macrophages, dendritic cells, NK cells, basophils, eosinophils, mast cells, and neutrophils. Whereas, lymphoid progenitor cells terminally differentiate into adaptive T/B immune cells[3,4].

Innate immunity evolved before the divergence of vertebrates and invertebrates and is critical for the survival of all multicellular organisms from plants to mammals[5]. While differences are present in innate immunity across organisms, the central mechanisms of defense are highly evolutionarily conserved. To recognize foreign pathogens, innate immune cells evolved pattern recognition receptors (PRRs). These germline-encoded receptors recognize conserved elements of pathogenic invaders. The first PRRs were discovered in drosophila for a role in defending against fungal infection [6–9]. Soon after the initial discovery of Toll-like receptors in drosophila, scientists found homologs in the human and mouse genomes[7]. To date, many PRRs have been described that include both intracellular sensors and cell membrane-bound receptors[10]. PRRs include Toll-

like receptors (TLRs), nucleotide-oligomerization domain (NOD)-like receptors, retinoic acid-inducible gene-I (RIG-I) receptors, C-type lectin receptors (CLRs), and absent in melanoma-2 (AIM2) receptors. PRRs bind to conserved pathogen-associated molecular patterns (PAMPs) such as flagellin, dsRNA, lipopolysaccharides (LPS), or unmethylated DNA[11,12]. For an extensive list of PRRs and the PAMPs that they recognize, see Table-1-1[13].

Innate immune cells rapidly respond to infection following the recognition of PAMPS by PRRs through the production of proinflammatory cytokines, chemokines, and effector molecules. Another well-studied role of certain innate immune cells is their ability to directly kill foreign pathogens through phagocytosis. Following phagocytosis and digestion of the foreign pathogen, innate immune cells termed antigen-presenting cells (APCs) display foreign peptides on their cell surface receptors to activate adaptive immune cells. While the innate immune system is critical for a proper response to infection, it alone is often not sufficient for the complete clearance of a pathogenic infection.

The adaptive immune system evolved in vertebrates for a more specific and robust response to infection. Unlike innate immune cells that have highly conserved PRRs, adaptive immune cells have T-cell receptors (TCR) and B-cell receptors (BCRs) that are highly specific and variable. TCRs/BCRs are generated from a few hundred germ-line encoded elements that undergo V(D)J recombination to create millions of different TCR/BCR possibilities[14]. Adaptive immune cells are predominantly activated in secondary lymphoid tissues including the lymph nodes, spleen, and Peyer's Patches. In the lymph nodes, B-cells survey for foreign antigens. Once antigens displayed on follicular dendric cells are recognized by a BCR, the B-cell transitions into an activated state by upregulating surface molecules, internalizing the antigen, and entering the G1 phase of the cell cycle[15]. After proliferation and clonal expansion, B-cells differentiate into plasma cells, germinal centers, or memory B-cells. These processes are tightly regulated to create a proper antibody response to infection.

T-cells are divided into two groups, cytotoxic CD8+ and helper CD4+ T-cells. When a CD8+ T-cell recognizes an antigen presented on an MHC I receptor, a cascade of events is initiated that causes the T-cell to proliferate and kill cells infected with and presenting the appropriate antigen. CD8+ T-cells express perforin and granzyme, which form holes and induce apoptosis, to kill cells infected with pathogens, respectively[16]. On the other hand, CD4+ T-cells, termed helper T-cells, recognize MHC II molecules, which are only present on antigen-presenting cells (APCs) of the innate immune system like dendric cells and macrophages. CD4+ T-cells need three signals to activate and differentiate: 1) TCR interaction with MHC class II; 2) co-stimulatory interactions between CD28: CD28L; and 3) cytokines. CD4+ T-cells can differentiate into Th1, Th2, Th17, T-regulatory, or T-follicular, based on the surrounding environment and cytokines present during activation. Each CD4+ T-cell type has a specialized function in pathogenic resolution. CD4+ T-cells also help in proper CD8+ T-cell activation and B-cell activation including class switching and affinity maturation[5,16–18]. The innate and adaptive immune systems work together in intricate ways in healthy individuals to resolve infection and injury.

This brief overview of the immune system highlights some of the important immune cell types and mechanisms of activation. However, our knowledge of the immune system continues to expand. In recent years, additional immune cell subtypes have been defined such as innate lymphoid cells and dendric cell types. As technology advances, immunologists continue to unravel the complexity of the immune system by defining cell types, cellular interactions, and mechanisms of action.

MACROPHAGES

Tissue-resident macrophages are an important component of the immune system. Since they are at the site of infection or injury, they are among the first to respond through phagocytosis, cytokine production, and antigen presentation. Phagocytosis is one mechanism used by macrophages to

directly kill invading pathogens. This process includes the fusion of the macrophage to the pathogen; the engulfment of the pathogen creating a phagosome inside the cell; the fusion of the phagosome with a lysosome that contains digestive enzymes; and the degradation of the pathogen[19]. Activated macrophages also produce regulatory molecules, proinflammatory cytokines, chemokines, and effector molecules. Proinflammatory cytokines have many functions including amplification of the inflammatory signal, recruitment of monocytes to the site of injury, and vasodilation. Finally, activated macrophages can travel through the lymphatic system to nearby lymph nodes. Once in the lymph node, macrophages present foreign antigens on their cell surface via MHC class II receptors to T-cells. This process is critical for the activation of the adaptive immune system[19–22]. It is important to note that dendritic cells share many of these functions including antigen presentation and proinflammatory gene activation. However, we use macrophages as a model system in our studies due to the historical use of macrophages to explore transcriptional regulation[23].

Macrophages respond to numerous ligands in various tissue contexts by inducing the expression of hundreds of genes. Molecular immunologists have made progress in understanding stimulus-specific responses by exploring the specificity of receptor activation, temporal activation of downstream adapters, and combinatorial transcription factor gene expression modulation[24]. When a macrophage senses an inflammatory signal, it must transmit this signal from the receptor to the nucleus to modulate gene expression. This process includes 1) receptor-ligand interactions; 2) signaling cascades; 3) sequence-specific transcription factors (TFs), proteins that directly bind to DNA to modulate gene expression; and 4) regulators of chromatin structure and post-transcriptional processes. Each step is fine-tuned for an appropriate response to inflammatory signals. While stimulus-specific PRRs begin to explain the differential activation of macrophages by various PAMPs, downstream signaling pathways can converge on the activation of conserved TFs. For example, one transcription factor that is activated downstream of all the TLRs is the NF-

κB family (Figure 1-1). While much progress has been made in understanding immune cell activation, a challenge in the field is defining the underlying mechanisms of macrophage specificity.

Differential activation of macrophages by tumor necrosis factor (TNF) and lipopolysaccharides (LPS) serves as an example of the stimulus-specific nature of inflammatory responses. TNF is a proinflammatory cytokine that can be secreted by macrophages, T-cells, NK cells, and some epithelial cells in response to microbial pathogens or cellular stress. In contrast, LPS is found on the outside of gram-negative bacteria. Deciphering how these two distinct stimuli lead to differential activation of macrophages has been of interest to molecular immunologists.

TNF was first described in the mid-1970s as a factor in endotoxin-treated serum for its ability to lyse tumor cells[25,26]. Today, TNF is well-characterized as a cytokine with roles in clearing pathogenic infection, inducing inflammation, and promoting apoptosis. TNF can signal through two membrane-bound receptors, TNFR1 and TNFR2. While TNFR1 is found on most cell types, TNFR2 is primarily found on immune cells. The activation of TNFR2 signals through TNFR-associated factor 1 (TRAF1) or TRAF2 to promote cell survival and activate TFs like NF-κB. The activation of TNFR1 leads to the aggregation of the intracellular domain that is recognized by the TNFR-associated death domain (TRADD)[27]. This cascade leads to the activation of transcription factors like NF-κB and MAPKs (Figure 1-2)[28–30]. However, depending on downstream signaling events and the cell context, activation of TNFR1 can either promote cell survival or death.

In contrast to TNF, which is a cytokine produced by endogenous cells in response to injury or infection, LPS is found on the outside of gram-negative bacteria. LPS is composed of lipid A, O-antigen, and hydrophilic core polysaccharides[31]. LPS binds to TLR4, which potently activates macrophages leading to transcriptome-wide changes. Hundreds of genes are upregulated in macrophages upon stimulation with LPS, many of which are pro-inflammatory cytokines (*Il1b, Tnf,*

Il6), type I interferons *(Ifnb*), and anti-microbial agents (*Nos2, Lcn2*)[32]. LPS binding to TLR4 activates two distinct pathways that signal through either myeloid differentiation primary response 88 (MyD88) or Toll/interleukin-1 receptor-domain-containing adapter-inducing interferon-beta (TRIF). The adapter MyD88 is activated rapidly through dimerization of the TLR4 domain upon LPS binding. MyD88 begins the signaling cascade that leads to the activation of transcription factors like NF-κB. Sequentially, TLR4 endocytosis activates TRIF, which leads to robust and sustained activation of both NF-κB and IRF3 (Figure 1-1)[33,34].

There is substantial interest in further understanding mechanisms that could explain transcriptional differences between macrophages activated with TNF and LPS. To date, two main mechanisms have been elucidated. First is a difference in transcription factor activation. Whereas LPS activates NF-κB, IRF3, AP1, and p38/CREB; TNF activates NF-κB, p38/CREB, and AP1 but not IRF3. Second is a difference in the temporal activation of NF-κB[35]. TNF stimulation leads to the rapid but transient activation of NF-κB, whereas LPS signaling leads to more robust and sustained activation of NF-κB. While these two mechanisms account for some differences between macrophages activated by TNF or LPS, fully understanding the distinct transcriptional states is far from complete.

TRANSCRIPTIONAL GENE REGULATION

All cells must integrate signals from the environment to properly modulate gene expression. Gene regulation occurs through multiple mechanisms, including transcriptional and post-transcriptional processes, such as mRNA splicing, mRNA stability, and translation. Since transcriptional regulation is the most prominent mechanism that affects gene expression, it is the focus of this study. The transcriptional regulation of genes is a complex process that involves DNA elements, regulatory proteins, and transcriptional machinery. Key DNA elements that influence transcription include promoter and enhancer regions. The promoter is found approximately +/- 150bps

surrounding the transcription start site (TSS). Important components of the core promoter region are the TATA box and initiator, with sequence-specific TF binding motifs generally located distal to the core promoter[36]. These DNA elements help recruit the transcription machinery and nucleate the formation of the pre-initiation complex (PIC). While the core promoter elements were originally thought to be conserved at all gene promoters, recent studies reveal more flexibility[37].

General transcription factors, including TFIIA, TFIIB, TFIID, TFIIF, TFIIS, TFIIE, and TFIIH, often assemble with RNA polymerase II to form the PIC. Once the PIC assembles, it initiates transcription by melting the DNA, inserting the single-stranded DNA into the Pol II active site, and catalyzing the transcription of DNA to RNA. Once the nascent RNA strand grows, the PIC is cleared from the promoter and the stable elongation complex forms. The elongation complex is stabilized by the interaction between RNA Pol II, the DNA template, and the nascent RNA transcript[37–41].

Sequence-specific transcription factors (TFs) help coordinate basal transcription, as well as transcriptional changes during development and in response to stimuli. They can directly influence gene expression through binding to the promoter/enhancer regions; recruitment of Mediator and/or other components of the general transcription machinery; as well as stabilization of the PIC. Enhancer regions can be up to mega-bases away from the TSS. CTCF and cohesin can affect chromatin architecture by creating topologically associated domains (TAD), which are loops in DNA that bring enhancer and promoter regions closer together (Figure 1-3)[42]. TFs can also bind regulatory regions to either directly alter the chromatin structure or recruit co-activating proteins that can modulate chromatin structure. Several different DNA sequence-specific transcription factors typically bind to a collection of DNA motifs that comprise a promoter or an enhancer.

THE NF-κB FAMILY

Since the discovery of the NF-κB family of transcription factors in 1986 by Ranjan Sen and David Baltimore, there has been extensive research investigating its structure and function[43]. In the past 37 years, the NF-κB family has proven to be a critical mediator of immune activation and a necessary player in proper immune cell development. The NF-κB family contains five subunits, c-Rel (*Rel)*, p65 (*Rela),* RelB (*Relb)*, p50 (*Nfkb1*), and p52 (*Nfkb2*), that can combinatorially combine to form 15 possible homo- and heterodimers[44,45]. Because all five NF-κB subunits are expressed in immune cells, there has been interest in understanding the immunological logic underlying the combinatorial complexity of this system; the distinct function of each dimer; and the underlying mechanisms that would allow different dimers of NF-κB to regulate distinct sets of genes.

The NF-κB family of transcription factors share a Rel homology region (RHR), a 300 amino acid N-terminal domain that includes a DNA binding domain, allowing the TF to directly interact with the major groove of DNA in a sequence-specific manner; a dimerization domain that facilitates dimerization with other NF-κB subunits; a scaffold for interaction with IκBs; and a nuclear localization sequence. RelA, c-Rel, and RelB also contain an activation domain, which allows an NF-κB dimer with one of these subunits to act as a transcriptional activator by creating a binding scaffold for other transcriptional coregulators. Conversely, p50 and p52 do not contain an activation domain. This means p50 homodimers, p52 homodimers, and p50:p52 heterodimers cannot act alone as transcriptional activators. Notably, p50 and p52 are transcribed and translated as larger precursor proteins, p105 and p100, respectively. These precursor proteins contain ankyrin repeats (30-34 amino acid repeated sequences), which block the nuclear localization sequence and aid in sequestering NF-κB in the cytosol. Thus, the production of functional p50 and p52 requires an additional post-translational cleavage of the ankyrin repeats[46]. While p105 is

constitutively processed to form functional p50; p100 processing is more tightly regulated during the non-canonical activation pathway of NF-κB [47–49].

In unstimulated cells, NF-κB genes are basally transcribed and translated into protein. Upon translation, NF-κB subunits immediately dimerize and are bound by inhibitors of NF-κB (IκBs), which prevents their nuclear localization. Two main signaling pathways lead to the activation of NF-κB, the canonical and non-canonical activation pathways. The canonical activation pathway is downstream of diverse receptor-ligand interactions including all the TLRs. This intracellular signaling cascade leads to the activation of TGFβ–activated kinase 1 (TAK1), which in turn activates the IKK complex. The IKK trimeric complex is composed of IKKα, IKKβ, and IKKγ. Once activated, the IKK complex auto-phosphorylates and then phosphorylates IκBs. After the IKK complex phosphorylates IκBs, they are ubiquitinated and tagged for proteasomal degradation. This releases NF-κB to translocate into the nucleus (Figure 1-4)[50–52]. In contrast, the non-canonical activation pathway is activated mainly by a subset of TNFR superfamily members, most notably BAFF and CD40. Upon activation, the non-canonical pathway signals through NF-κB-inducing kinase (NIK) that activates IKKα and leads to the phosphorylation, ubiquitination, and proteasomal processing of p100 into p52. This allows p52-containing dimers, mainly p52 homodimers and p52:RelB heterodimers, to translocate into the nucleus. Compared to the rapid and transient activation of NF-κB through the canonical activation pathway, the non-canonical pathway is slower and more prolonged (Figure 1-4)[53–55].

NF-κB SELECTIVE FUNCTION

While NF-κB is one of the best-studied transcription factors, with almost 100,000 publications cited on Pub-Med, there is still a paucity of information describing the underlying molecular mechanisms of NF-κB selective functions. In the late 1990s, mice lacking *Rel*, *Nfkb1*, *Nfkb2*, *Relb*, and *Rela* were generated using homologous recombination to disrupt the gene of interest in

embryonic stem cells (ESCs). While mice lacking *Rel*, *Relb*, *Nfkb1*, and *Nfkb2* are viable, mice lacking *Rela* are embryonic lethal[56]. However, mice heterozygous for *Rela* are viable and develop normally. Mice with homozygous deletion of *Rela* do not live past day 15-16 in gestation due to TNFα-induced cell death of hepatocytes which leads to liver degradation. This makes it difficult to study the effect of loss of *Rela* on the immune system, since at day 15-16 in gestation the hematopoietic precursors are intact, but the immune system is still developing.

The generation of mice lacking *Nfkb1* reveals that while *Nfkb1* is not critical for development, it is important in B-cell responses and the resolution of infection. *Nfkb1*$^{-/-}$ mice show significant defects in B-cell proliferation in response to LPS; antibody production; and resolution of L. monocytogenes or S. pneumoniae infections. However, *Nfkb1*$^{-/-}$ mice show an enhanced ability to clear murine encephalomyocarditis virus. These data indicate that *Nfkb1* plays a multifaceted and non-redundant role in proper immune system activation[57–59].

Mice lacking *Relb* have severe immunological defects that can lead to premature death around day 20. Immunological defects observed in *Relb*$^{-/-}$ mice include splenomegaly; a decrease in thymic dendritic cells; myeloid hyperplasia in the bone marrow; lymphoid depletion of lymph nodes; and an inflamed lung and liver. These data suggest that *Relb* is critical for proper hematopoiesis, which cannot be compensated for by other members of the NF-κB family[60].

Loss of *Nfkb2* leads to defects in the spleen and lymph node architecture and the inability to form germinal centers. The inability of mice to mount proper antibody responses is thought to be more a function of defects in antigen presentation than intrinsic B-cell properties since T/B cell proliferation is only mildly affected by the loss of *Nfkb2* when activated artificially in cell culture[61].

Phenotypic studies indicate that c-Rel has a distinct function in both innate and adaptive immunity. Although the overall number of hematopoietic cells appears normal, various immune cells,

including B-cells, T-cells, and macrophages, exhibit impaired activation or differentiation in *Rel*$^{-/-}$ mice. B-cell defects include impaired isotype switching and proliferation due to cell cycle arrest and increased apoptosis. T-cells exhibit defects in overall proliferation; T-reg differentiation; and cytokine production including IL-2, IL-3, and GM-CSF. There are also deficiencies observed in innate immunity, such as a lack of *Il12b*, *Il6*, and *Tnf* expression in peritoneal macrophages[62,63].

These phenotypic studies together illuminate the non-redundant functions of individual NF-κB subunits in immune cell development and/or activation. For a comprehensive list of NF-κB transgenic mice and their observed phenotypes, see Table 1-2. While these studies reveal immunological deficiencies, it is still unclear which target genes are responsible for observed immunological defects; if distinct dimers of NF-κB regulate unique sets of genes; and if so, what are the underlying mechanisms that allow different NF-κB dimers to regulate distinct genes.

MECHANISMS of NF-κB GENE REGULATION

Molecular biologists have investigated various aspects of the NF-κB system in order to gain insight into its selective function. They have examined the NF-κB repertoire in various cell types, studied the structure of NF-κB dimers bound to DNA, and examined the biochemical properties that distinguish NF-κB dimers. These studies aim to uncover the underlying mechanisms of NF-κB selectivity.

The NF-κB repertoire can vary in a cell-type-specific manner. For instance, while murine embryonic fibroblasts (MEFs) express an abundance of RelA:p50 heterodimers, RelA homodimers, and p50 homodimers, other dimers are scarce, including c-Rel- and RelB-containing dimers[64]. Another study demonstrated shifts in the NF-κB repertoire in differentiating B-cells. While B-cell precursors express an abundance of p50- or RelA-containing dimers, mature B-cells show a shift to c-Rel- or p50-containing dimers, and terminally differentiated plasma cells

primarily express p52- or RelB-containing dimers[65]. These studies reveal the NF-κB repertoire can be cell-type specific and may aid in cell differentiation.

NF-κB dimer formation is influenced not only by the cell context but also by the intrinsic properties of the NF-κB monomers and rates of protein synthesis. Since NF-κB monomers are rapidly degraded, the stability and affinity of dimers can shift the overall NF-κB repertoire[66]. For example, a co-refolding experiment revealed that c-Rel homodimers are more stable than RelA homodimers, consequently reducing their interactions with p50. Furthermore, rates of protein synthesis will also affect relative dimer abundance. While all NF-κB members are transcribed basally in macrophages, RelA-containing dimers control the stimuli-activated transcription of *Rel*, *Relb*, *Nfkb1*, and *Nfkb2*. While c-Rel, RelA, and RelB are translated as final monomeric proteins; p105 and p100 must undergo proteolytic cleavage to generate p50 and p52, respectively. Together, many factors influence NF-κB dimer formation including the cell context, dimerization stability, and protein synthesis[67].

Most immune cells, including macrophages, express all five NF-κB subunits. This makes it possible for all 15 NF-κB dimers to be present in the same cell. While all 15 dimers have the potential to form, three of these complexes (RelA:RelB, c-Rel:RelB, RelB:RelB) have not been shown to bind to DNA. This leaves 12 dimers that can potentially modulate gene expression.

Understanding the select function of unique NF-κB dimers has been of great interest to biologists. In the early 1990s, the 3-dimensional folding of NF-κB was revealed by crystal structures of NF-κB dimers bound to κB DNA sites. The x-ray crystal structure of a p50 homodimer bound to DNA was the first to be resolved to 2.3Å. This structure revealed that the RHR folds into two distinct domains, which are connected by a 10 amino acid linker sequence. Dimers are held together by β–sheets, where β-strands from each subunit's dimerization domain come together. Loops from

the N-terminal and C-terminal domains mediate DNA contacts[68]. In 1998, the RelA:p50 heterodimer crystal structure revealed interesting insights into the underlying mechanisms of dimer-specific DNA-binding and function. This crystal structure revealed that the direct protein-DNA interactions of p50 and RelA differ. While RelA can interact with the 4-bp sequence TTCC; p50 can directly interact with the 5-bp sequence GGGAC[69]. This specificity implies different dimers of NF-κB may have varying affinities to unique NF-κB motifs.

Of particular interest is understanding how c-Rel and RelA control distinct biological functions but have highly similar biochemical properties. As described earlier, *Rela*$^{-/-}$ exhibits an embryonic lethal phenotype, whereas *Rel*$^{-/-}$ mice develop normally. Furthermore, the loss of c-Rel leads to unique defects in both innate and adaptive immune cell activation. *Rela* is ubiquitously expressed, while *Rel* is almost exclusively expressed in immune cells along with *Rela* and *Nfkb1*. However, loss of *Rel* leads to interesting immunological phenotypes including CD4+ T-cells deficiency in IL-2 and GM-CSF production; defects in T-reg differentiation; and the loss of pDC/macrophage production of IL-12[70-71]. Both c-Rel and RelA are often present in the same cell types; are activated by the canonical activation pathway; contain activation domains; and have identical DNA binding contact residues for both the DNA backbone and within κB sites. This begs the question: What intrinsic properties of c-Rel and RelA can explain observed immunological differences? One study illustrated that in CD4+ T-cells, c-Rel is activated through CD28 costimulatory signaling. Following the activation and nuclear translocation, cRel-containing dimers bind to DNA elements coined the CD28 response elements (RE). The IL-2 CD28 RE contains a non-consensus NF-κB motif 5'-AGAAATTCC'-3 which appears to be bound by c-Rel homodimers. This study began to explore mechanisms underlying the distinct regulation of gene expression by different NF-κB dimers[72,73].

Another study explored the unique function of c-Rel in macrophage activation. Using c-Rel and RelA chimeric proteins, 46 residues in the c-Rel RHR were identified that are both necessary and

sufficient for the production of IL-12 in macrophages stimulated with LPS. This study further revealed that the 46 residue segment influences binding affinity to non-consensus NF-κB motifs[74]. Thus, c-Rel is indispensable for the expression on *Il12b* since non-consensus NF-κB motifs are exclusively found at the *Il12b* promoter region. These studies elucidated mechanisms by which transcription factors belonging to the same family can differentially regulate specific sets of genes in response to identical stimuli and cellular environments.

Another interesting example of NF-κB dimer-specificity has been described when exploring p50 homodimers. Since p50 does not contain an activation domain, p50 homodimers have been thought to act as gene repressors either by competing with other activation domain-containing NF-κB dimers or through the recruitment of histone deacetylases or histone methyltransferases[75,76]. Unlike other NF-κB complexes, p50 homodimers are present in the nucleus of unstimulated macrophages[77]. Thus, p50 homodimers have been implicated in repressing gene expression of inflammatory genes in unstimulated cells by directly blocking cis-regulatory elements. Additionally, p50 homodimers have been shown to repress the expression of interferon response genes in macrophages stimulated with LPS by binding to guanine-rich sequences at or near the interferon response element[78].

Other studies revealed an interaction between p50 homodimers and proteins outside of the NF-κB family, such as nuclear IκBs. Nuclear IκBs have been shown to form ternary complexes with p50 homodimers and the kB site to activate gene expression[79–81]. Nuclear IκBs are a family of proteins, including IκBζ, Blc3, and IκBNS, which are closely related to cytosolic IκBs. Unlike cytosolic IκBs that have a known role of sequestering NF-κB in the cytosol, nuclear IκBs are induced upon stimulation and translocate into the nucleus. Although nuclear IκBs are not well classified, co-immunoprecipitations show that p50 selectively interacts with IκBζ to regulate pro-inflammatory genes like *Il6, Il12b,* and *Csf1*[82]*.* While these studies provided evidence that p50

homodimers associate with proteins outside of the NF-κB family to activate gene expression, the underlying mechanisms of regulation are still unclear.

Exploring sequence-specific preferences of individual NF-κB dimers can help explain the ability of distinct dimers to regulate unique sets of genes. The first extensive unbiased study exploring preferential binding by eight different NF-κB dimers was carried out using surface plasmon resonance and protein binding microarrays. These studies revealed three distinct NF-κB dimer motifs that specify either 1) p50 or p52 homodimers, 2) heterodimers, or 3) c-Rel or RelA homodimers. Furthermore, affinity differences were also observed. In comparison to other dimers, c-Rel homodimers demonstrate the highest affinity to non-consensus NF-κB motifs. These findings reveal valuable insights into the intrinsic properties of NF-κB dimers that underlie their selective functions[83]. Overall, while much has been learned about NF-κB, we are just beginning to understand the selective roles of NF-κB in immune cell activation.

THE IκB FAMILY

IκBs play a crucial role in the regulation of NF-κB activity by increasing dimerization stability and hindering nuclear translocation[84]. The IκB family can be subdivided into cytoplasmic and nuclear IκBs. Cytoplasmic IκBs (IκBα, IκBβ, IκBε) are known to sequester NF-κB in the cytosol in unstimulated cells by blocking the nuclear localization sequence. Upon stimulation, cytoplasmic IκBs are phosphorylated, ubiquitinated, and tagged for proteasomal degradation, which leads to the release of NF-κB for nuclear translocation. NF-κB directly targets and induces the expression of cytosolic IκBs, which generates a negative feedback loop for NF-κB regulation. Notably, IκBα has been shown to translocate into the nucleus to bind NF-κB, thereby drawing it out of the nucleus and shutting down the activation of NF-κB target genes[85]. IκBβ inserted under the promoter of IκBα can compensate for the loss of IκBα. However, mice lacking IκBβ or IκBα have

distinct immunological defects, suggesting these proteins have unique endogenous biological functions.

Nuclear IκBs (IκBζ, Bcl3, and IκBNS) are induced upon stimulation and translocate to the nucleus[86]. While nuclear IκBs have been shown to interact with NF-κB, their specific immunological roles and mechanistic functions remain elusive. Of particular interest are IκBζ (*Nfkbiz*) and Bcl3 (*Bcl3*) since both contain activation domains and may associate with NF-κB to facilitate gene activation[87]. However, we do not see a role for Blc3 in macrophage activation.

Studies were performed to understand which inflammatory signals directly activate IκBζ in macrophages. Although IκBζ is induced in response to TLR1/2/4/5/6/7/9 signaling, it is not induced by TNFα. Moreover, LPS induction of *Nfkbiz* is abolished in *Myd88*$^{-/-}$ mice, suggesting that *Nfkbiz* expression is dependent on the MyD88 pathway in TLR signaling[88]. IκBζ is regulated both at the expression level and post-transcriptionally. Differences in *Nfkbiz* expression between TNF and LPS from mRNA data sets are substantial. However, differences in chromatin-associated RNA expression of *Nfkbiz* are less dramatic, indicating that *Nfkbiz* is additionally regulated through mRNA stability and degradation[89].

In 2004, *Nfkbiz*$^{-/-}$ mice were generated through targeted disruption of the *Nfkbiz* gene locus with the insertion of a Neo cassette into ESCs[82]. Although viable, *Nfkbiz*$^{-/-}$ mice develop atopic dermatitis around 4-5 weeks of age; exhibit defects in splenocyte proliferation in response to LPS; have reduced numbers of CD4+ T-cells; and demonstrate increased lymphocyte infiltration into the submucosa, the conjunctival epithelium, and the lungs. Additionally, macrophages from *Nfkbiz*$^{-/-}$ mice have defects in IL-6 and IL-12 production in response to LPS. These studies indicate that IκBζ plays a unique and interesting role in proper immune cell activation.

ADDRESSING NF-κB SPECIFICITY

The study of NF-κB specificity has been hindered by significant challenges that arise from redundancy among NF-κB subunits; the combinatorial complexity of multiple dimers; and the similarity in DNA recognition motifs. Although the deletion of individual NF-κB subunits has uncovered an interesting role for each subunit in immune cell activation, the direct gene targets of NF-κB dimers and the mechanisms underlying regulation remain largely unknown. Moreover, detangling subunit function from dimer function is complicated, as the loss of a single NF-κB subunit can abolish up to five dimers. For instance, the deletion of *Nfkb1* in macrophages leads to the loss of p50 homodimers, p50:p52, p50:RelA, p50:RelB, and p50:52 heterodimers. Thus, unraveling dimer-specific functions has been a great challenge for this field.

The emergence of genomic techniques, such as RNA sequencing (RNA-seq) and chromatin immunoprecipitation sequencing (ChIP-seq), presents an opportunity to explore NF-κB specificity in thorough, unbiased, and quantitative ways. Our lab employs bone marrow-derived macrophages (BMDMs) stimulated with LPS or lipid A as a model system to study NF-κB dimer-specific function. The use of genomics approaches in a reductionist system allows for the unprecedented study of NF-κB in immune cell activation.

The use of chromatin-associated RNA-seq in macrophages lacking a subunit of NF-κB, provides insight into transcriptome-wide gene dependencies during macrophage activation. The direct role of NF-κB is explored by capturing changes in the transcriptome immediately after stimulation (30min – 120 minutes). This time frame enables the study of the direct effects of NF-κB and excludes secondary effects that may accumulate at later time points. After defining subunit function through RNA-seq, we begin exploring dimer-specific binding and its role in transcriptional regulation by employing ChIP-seq. Integration of ChIP-seq datasets from all five NF-κB subunits provides an initial understanding of binding preferences and dimer complex interactions genome-

wide. Sequential ChIP-seq allows for the direct exploration of subunit interaction in the cell context. Sequential ChIP-seq works by first immunoprecipitating a protein of interest, and then using a second antibody to sequentially immunoprecipitate a potential interaction partner of the first protein. The resulting DNA reveals locations in the genome that are directly bound by both proteins of interest.

By using the genomics approaches described, in combination with *in vitro* assays such as co-immunoprecipitations and EMSAs, we explore NF-κB dimer-specific function with unprecedented depth. Chapter 2 describes how ChIP-seq can be used to capture and understand NF-κB selective interactions in the cell context. In Chapter 3, we explore the role of p50 in macrophages activated with LPS. We define the non-redundant roles of p50 in modulating gene expression; investigate underlying mechanisms of p50-dependent gene expression; and explore the immunoregulatory significance of p50 selectivity. Chapter 4 delves into the unique role of c-Rel in macrophage activation. We highlight the select role of cRel in activating *Il12b,* a cytokine involved in T-cell differentiation. We demonstrate that the evolutionary divergence of NF-κB subunits coincides with the emergence of the adaptive immune system in vertebrates. Together these findings illuminate NF-κB selectivity in immune cell activation and define novel mechanisms that explain NF-κB selective functions.

FIGURE LEGENDS

Figure 1-1: Activation of Innate Immune Cells by TLRs

Macrophages are activated by various TLRs that recognize conserved pathogen-associated molecular patterns. All TLRs activate the NF-κB family of transcription factors. TLR4 is activated by LPS and signals through adapter MyD88 and TRIF which lead to the nuclear translocation of TFs including NF-κB, CREB, AP1, and IRF3.

Figure 1-2: The TNF Signaling Pathways

TNF signals through TNF-receptors, which activate the adapters TRADD, TRAF1, or TRAF2 that begin the intracellular signaling cascade. The signaling cascade culminates in the activation of 1) TAK1 which is an upstream activator of p38/CREB and 2) IKKα, IKKβ, and NEMO, which in turn activate NF-κB.

Figure 1-3: 3-D Chromatin Structure and Transcriptional Gene Regulation

Topologically associated domains (TADs) are loops formed in chromatin by CTCF and Cohesin. TADs enable interactions between promoter and enhancer regions that are located kilobases to mega-bases apart. Mediator can interact with sequence-specific TFs to recruit cohesin, which can transiently support further looping within the TAD. Interactions between distal TFs, mediator, and the pre-initiation complex (PIC) can stabilize the PIC and enhance transcription.

Figure 1-4: Canonical and Non-Canonical Activation Pathways of NF-κB

The canonical activation pathway for NF-κB begins with the activation of the IKK complex. Once the IKK complex is activated, it phosphorylates IκBs bound to NF-κB. After phosphorylation, IκBs are ubiquitinated and tagged for proteasomal degradation which releases NF-κB for nuclear translocation. In contrast, the non-canonical or alternative activation pathway signals through NIK to activate IKKα, which in turn phosphorylate p100. p100 is then processed to form functional p52. This releases p52 containing dimers for nuclear translocation.

Table 1-1: PRRs and the PAMPs They Recognize

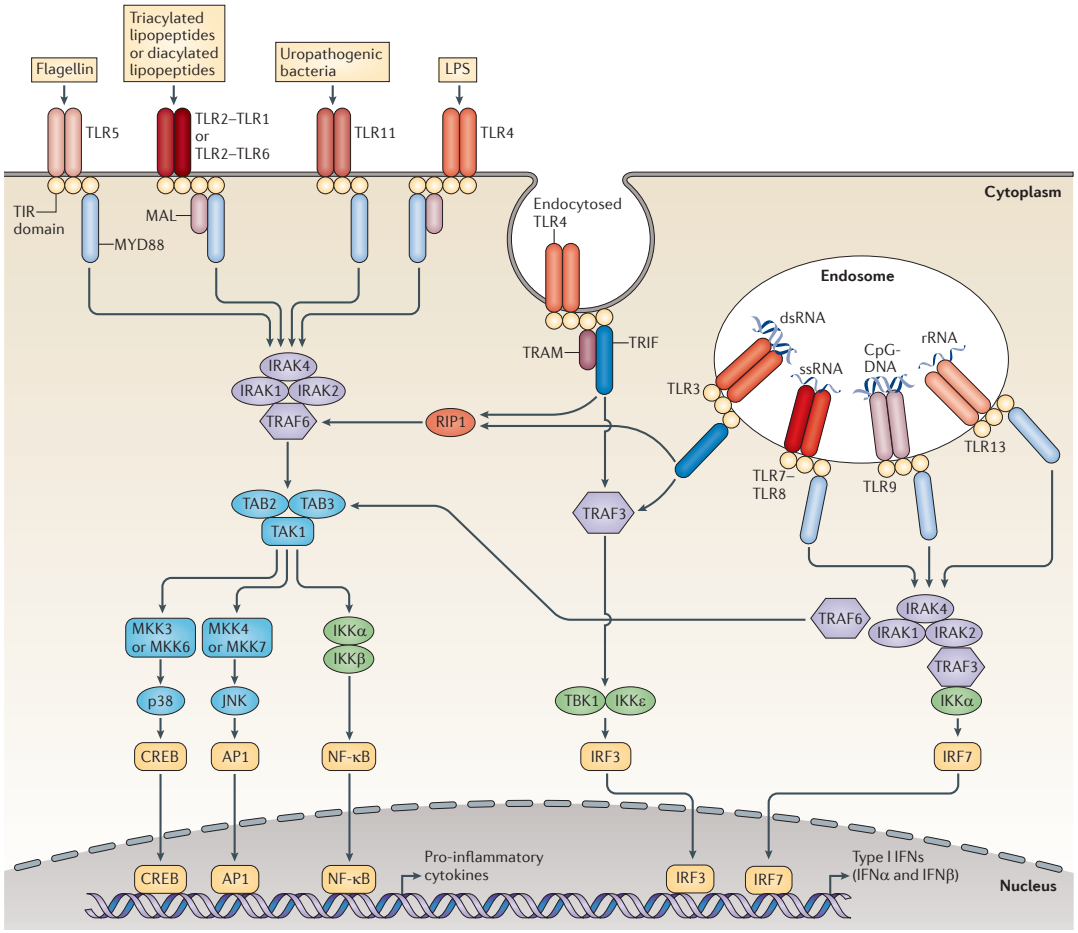
This table is a comprehensive list of pattern recognition receptors (PRRs) and the pathogen-associated molecular patterns (PAMPs) they recognize. PRRs can be both cell membrane-bound

receptors or cytoplasmic sensors, allowing innate immune cells to recognize extracellular infections, phagocytosed pathogens, and intracellular infections.

Table 1-2: NF-κB Transgenic Mice and Biological Phenotypes

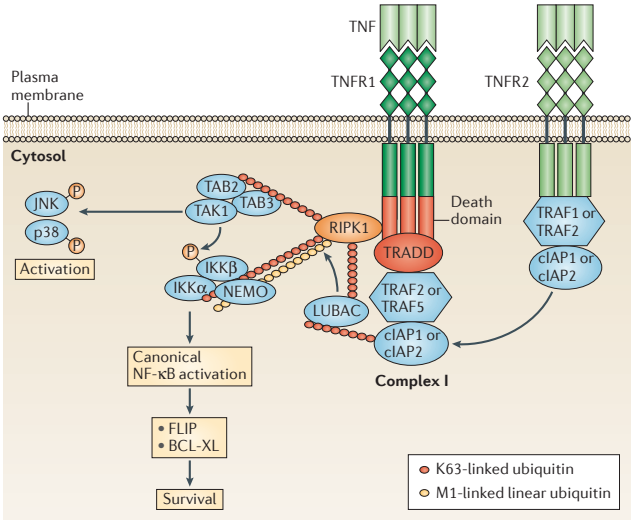
Mice lacking subunits of NF-κB have distinct biological and immunological phenotypes. All five subunits have been mutated in mice models to understand the role of each subunit in mouse development and immune cell activation. Further, mice with a C-terminal deletion of *Nfkb1* or *Nfkb2* lead to mice that lack the p100/p105 precursor but still contain functional p50/p52. Finally, double knock-out mice are generated to further explore the effects of NF-κB subunit redundancy.

Figure 1-1: Activation of Innate Immune Cells by TLRs



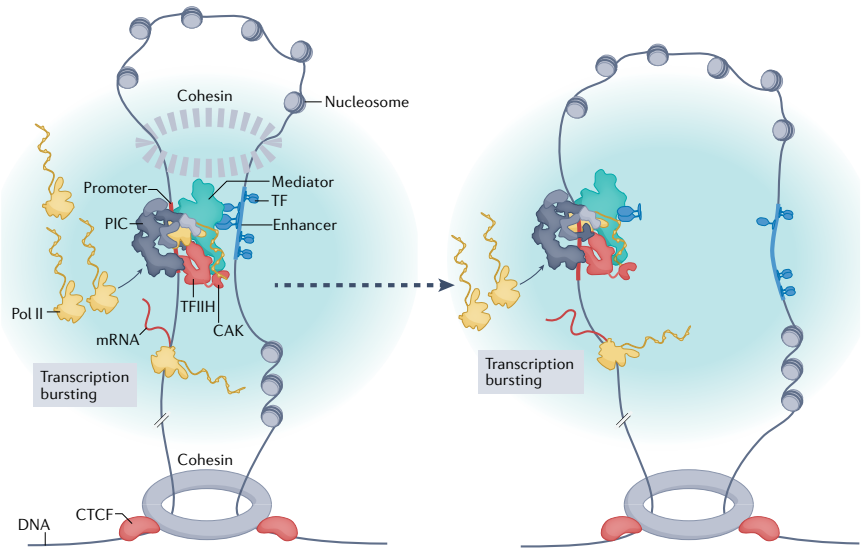
(O'Neil *et al.*, 2013)

Figure 1-2: The TNF Signaling Pathways



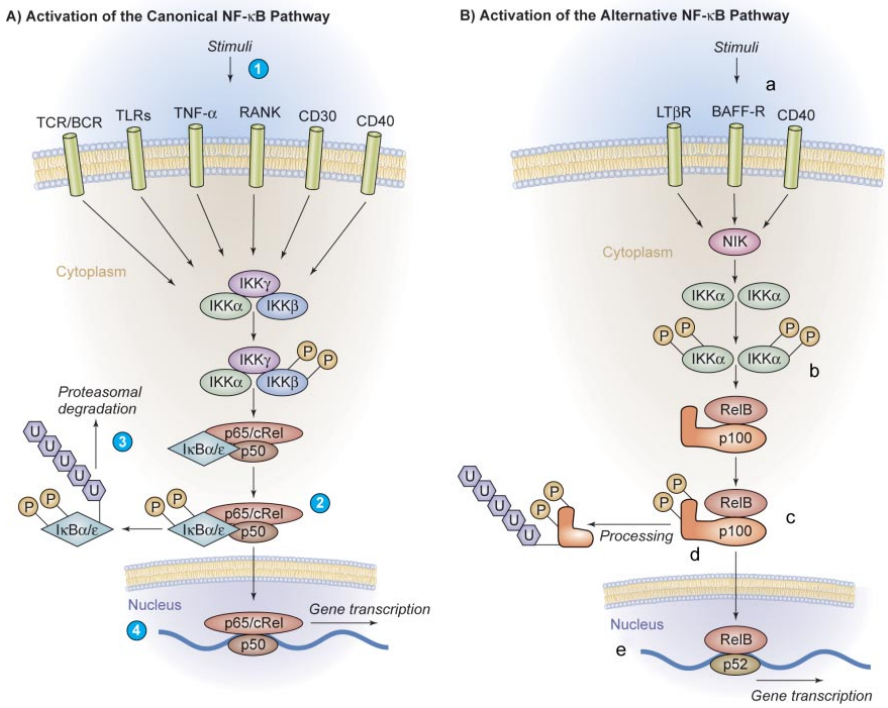
(Brenner, Blaser, and Mak, 2015)

Figure 1-3: 3-D Chromatin Structure and Transcriptional Gene Regulation



(Richter *et al.,* 2022)

Figure 1-4: Canonical and Non-Canonical Activation Pathways of NF-κB



(Jost and Ruland, 2007)

Table 1-1: PRRs and the PAMPs They Recognize

Table 1. PRRs and Their Ligands

PRRs	Localization	Ligand	Origin of the Ligand
TLR			
TLR1	Plasma membrane	Triacyl lipoprotein	Bacteria
TLR2	Plasma membrane	Lipoprotein	Bacteria, viruses, parasites, self
TLR3	Endolysosome	dsRNA	Virus
TLR4	Plasma membrane	LPS	Bacteria, viruses, self
TLR5	Plasma membrane	Flagellin	Bacteria
TLR6	Plasma membrane	Diacyl lipoprotein	Bacteria, viruses
TLR7 (human TLR8)	Endolysosome	ssRNA	Virus, bacteria, self
TLR9	Endolysosome	CpG-DNA	Virus, bacteria, protozoa, self
TLR10	Endolysosome	Unknown	Unknown
TLR11	Plasma membrane	Profilin-like molecule	Protozoa
RLR			
RIG-I	Cytoplasm	Short dsRNA, 5′triphosphate dsRNA	RNA viruses, DNA virus
MDA5	Cytoplasm	Long dsRNA	RNA viruses (Picornaviridae)
LGP2	Cytoplasm	Unknown	RNA viruses
NLR			
NOD1	Cytoplasm	iE-DAP	Bacteria
NOD2	Cytoplasm	MDP	Bacteria
CLR			
Dectin-1	Plasma membrane	β-Glucan	Fungi
Dectin-2	Plasma membrane	β-Glucan	Fungi
MINCLE	Plasma membrane	SAP130	Self, fungi

(Takeuchi and Akira, 2010)

Table 1-2: NF-κB Transgenic Mice and Biological Phenotypes

Genotype	*Lethality*	*Defect/phenotype*	*Reference*
Knockouts			
nfkb1 $^{-/-}$	No	B cells: marginal zone and CD5^{+} peritoneal B cells reduced; response to LPS diminished, turn over rapidly *in vivo*; defective isotype switching and impaired humoral immune response	Sha et al. (1995), Grumont et al. (1998), Snapper et al. (1996), Cariappa et al. (2000), Pohl et al. (2002)
		Th2 differentiation is impaired	Das et al. (2001), Artis et al. (2005)
		NK cells display enhanced proliferation and increased IFN*γ* production	Tato et al. (2006)
		Macrophages: ERK mitogen-activated protein kinase pathway activation in response to TLR signals is impaired leading to reduced expression of IL-6, IL-10 and Cox-2	Waterfield et al. (2003), Banerjee et al. (2006)
nfkb2 $^{-/-}$	No	Defective secondary lymphoid organ development, impaired B-cell development; enhanced DC function	Caamano et al. (1998), Franzoso et al. (1998), Speirs et al. (2004)
c-rel $^{-/-}$	No	B cells: cell-cycle and survival defects; impaired isotype switching	Grumont et al. (1998, 1999, Cheng et al. (2003), Pohl et al. (2002)
		T cells: defects in CD4 and CD8 T-cell responses, Th1 development and cytokine production (IL-2 and GM-CSF) in CD4^{+} T-cell responses	Hilliard et al. (2002), Lamhamedi-Cherradi et al. (2003), Mason et al. (2004), Gerondakis et al. (1996), Rao et al. (2003)
		Reduced number of pDC, impaired IL-12 production and DC priming of CTL	O'Keeffe et al. (2005), Grumont et al. (2001), Mintern et al. (2002)
		Neuronal survival defects	Pizzi et al. (2002, 2005)
rela $^{-/-}$	Yes (~E15)	TNF-*α*-induced cell death: hepatocytes, macrophages and fibroblasts	Beg et al. (1995b), Beg and Baltimore (1996), Prendes et al. (2003), Senftleben et al. (2001), Li et al. (2001)
		Impaired secondary lymphoid organ development	Alcamo et al. (2002)
		Defects in leukocyte recruitment, T-cell-dependent responses, isotype switching to IgG3	Alcamo et al. (2001, 2002, Horwitz et al. (1999)
		Spatial learning responses	Meffert et al. (2003)
		Epidermal homeostasis	Zhang et al. (2004)
relb $^{-/-}$	No	Complex inflammatory phenotype and hematopoietic abnormalities	Burkly et al. (1995), Weih et al. (1995)
		Defects in secondary lymphoid organ structure and germinal center formation	Weih et al. (2001), Yilmaz et al. (2003)
		Lack certain DC populations, DC functional defects	Wu et al. (1998), Zanetti et al. (2003), Castiglioni et al. (2002)
		Generation of early NK1.1^{-} NKT cells is impaired	Elewaut et al. (2003), Sivakumar et al. (2003)
		T cells: reduced proliferation and survival of single positive thymocytes; impaired Th1 differentiation	Guerin et al. (2002), Corn et al. (2005), Weih et al. (1997)
nfkb1 $^{\Delta CT/\Delta CT}$	No	Splenomegaly, enlarged lymph nodes and lymphoid infiltrates in various organs. Increase in B-cell numbers accompanied by hyper-responsiveness to mitogens. T cells exhibit diminished proliferative capacity	Ishikawa et al. (1998)
c-rel $^{\Delta CT/\Delta CT}$	No	Hypoplastic bone marrow, enlarged lymph node and lymphoid hyperplasia	Carrasco et al. (1998)
nfkb2 $^{\Delta CT/\Delta CT}$	No	Gastric hyperplasia, hyperkeratosis in the heart, spleen and thymic atrophy, enlarged lymph nodes, lymphocytic infiltrates of various tissues and granulocytosis	Ishikawa et al. (1997)

Mice lacking multiple subunits			
nfkb1 $^{-/-}$ *nfkb2* $^{-/-}$	No	Growth retarded and craniofacial abnormalities due to osteopetrosis	Franzoso et al. (1997b), Iotsova et al. (1997)
		Severe defects in secondary lymphoid organ development	Lo et al. (2006)
		Lack marginal zone and mature follicular B cells	Franzoso et al. (1997b)
nfkb1 $^{-/-}$ *relb* $^{-/-}$	No	Exacerbation of organ inflammation seen in *relb*$^{-/-}$ mice	Weih et al. (1997)
nfkb1 $^{-/-}$ *c-rel* $^{-/-}$	No	Immune defects more severe than individual mutants. Markedly diminished CD5^{+} peritoneal B cells	Pohl et al. (2002)
		B cells fail to undergo blast formation	Grumont et al. (2002)
		Impaired antigen-induced CD4^{+} T-cell responses	Zheng et al. (2003)
		Conventional and pDC markedly reduced, partly owing to survival defects; maturation of pDC in response to specific TLR signals is impaired	O'Keeffe et al. (2005)
nfkb1 $^{-/-}$ *rela* $^{-/-}$	Yes (~E13)	Defects in myelopoiesis and B-cell development in radiation chimeras	Horwitz et al. (1997)
rela $^{-/-}$ *c-rel* $^{-/-}$	Yes (~E13)	Multiple hemopoietic defects. Radiation chimeras exhibit nucleated erythrocytes, reduced number of B cells, systemic expansion of granulocytes and a reduction in monocytes	Grossmann et al. (1999, 2000)
		T cells exhibit a cell-cycle block early in G1 resulting from a failure to undergo c-Myc-dependent growth	Grumont et al. (2004)
rela $^{-/-}$ *c-rel* $^{-/-}$ *tnf* $^{-/-}$	Neonatal	E18 embryos exhibit multiple epidermal defects that include a failure to form specific hair types, disorganized basal keratinocyte architecture and thinner skin. RelA and c-Rel control skin homeostasis in new-born mice by suppressing the recruitment and activation of cells of the immune system	Gugasyan et al. (2004)

(Gerondakis *et al.*, 2006)

WORKS CITED

1. Laird, D. J., De Tomaso, A. W., Cooper, M. D. & Weissman, I. L. 50 million years of chordate evolution: Seeking the origins of adaptive immunity. Proc. Natl. Acad. Sci. U. S. A. 97, 6924–6926 (2000).
2. Janeway, C.A., Travers, P, Walport, M, et al. The components of the immune system. Immunobiol. Immune Syst. Heal. Dis. 5th Ed. New York Garl. Sci. (2001).
3. Pinho, S. & Frenette, P. S. Haematopoietic stem cell activity and interactions with the niche. Nat. Rev. Mol. Cell Biol. 20, 303–320 (2019).
4. Majeti, R., Park, C. Y. & Weissman, I. L. Identification of a hierarchy of multipotent hematopoietic progenitors in human cord blood. Cell Stem Cell 1, 635–645 (2007).
5. Cooper, M. D. & Alder, M. N. The evolution of adaptive immune systems. Cell 124, 815–822 (2006).
6. Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M. & Hoffmann, J. A. The dorsoventral regulatory gene cassette spatzle/Toll/Cactus controls the potent antifungal response in Drosophila adults. Cell 86, 973–983 (1996).
7. Medzhitov, R., Preston-Hurlburt, P. & Janeway, C. A. A human homologue of the Drosophila toll protein signals activation of adaptive immunity. Nature 388, 394–397 (1997).
8. Gay, N.J. & Keith, F. J. Drosophila Toll and IL·1 receptor. Nature 351, 355–356 (1991).
9. Hashimoto, C., Hudson, K.L. & Anderson, K. The Toll gene of Drosophila, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. Cell 52, 269–79 (1988).
10. Li, D. & Wu, M. Pattern recognition receptors in health and diseases. Signal Transduct. Target. Ther. 6, 1–24 (2021).

11. O'Neill, L. A. J., Golenbock, D. & Bowie, A. G. The history of Toll-like receptors-redefining innate immunity. Nat. Rev. Immunol. 13, 453–460 (2013).
12. Paludan, S. R., Pradeu, T., Masters, S. L. & Mogensen, T. H. Constitutive immune mechanisms: mediators of host defence and immune regulation. Nat. Rev. Immunol. 21, 137–150 (2021).
13. Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. Cell 140, 805–820 (2010).
14. Schatz, D. G. & Ji, Y. Recombination centres and the orchestration of V(D)J recombination. Nat. Rev. Immunol. 2011 114 11, 251–263 (2011).
15. Cyster, J. G. & Allen, C. D. C. B-Cell responses: cell interaction dynamics and decisions. Cell 177, 524–540 (2019).
16. Bevan, M. J. Helping the CD8+ T-cell response. Nat. Rev. Immunol. 4, 595–602 (2004).
17. Borst, J., Ahrends, T., Bąbała, N., Melief, C. J. M. & Kastenmüller, W. CD4+ T cell help in cancer immunology and immunotherapy. Nat. Rev. Immunol. 18, 635–647 (2018).
18. Laidlaw, B. J., Craft, J. E. & Kaech, S. M. The multifaceted role of CD4+ T cells in CD8+ T cell memory. Nat. Rev. Immunol. 16, 102–111 (2016).
19. Underhill, D. M. & Goodridge, H. S. Myeloid phagocytes-including macrophages, dendritic cells (DCs) and neutrophils-have numerous roles in immunity. (2012). doi:10.1038/nri3244
20. Murray, P. J. et al. Macrophage activation and polarization: nomenclature and experimental guidelines. Immunity 41, 14–20 (2014).
21. Mass, E., Nimmerjahn, F., Kierdorf, K. & Schlitzer, A. Tissue-specific macrophages: how they develop and choreograph tissue biology. Nat Rev Immunol (2023). doi:10.1038/s41577-023-00848-y
22. Mosser, D. M., Hamidzadeh, K. & Goncalves, R. Macrophages and the maintenance of homeostasis. Cell. Mol. Immunol. 2020 183 18, 579–587 (2020).

23. Smale, S. T. Hierarchies of NF-κB target-gene regulation. Nat. Immunol. 2011 128 12, 689–694 (2011).

24. Werner, S. L., Barken, D. & Hoffmann, A. Stimulus specificity of gene expression programs determined by temporal control of IKK activity. Science (80-.). 309, 1857–1861 (2005).

25. Carswell, E. A. et al. An endotoxin-induced serum factor that causes necrosis of tumors. Proc. Natl. Acad. Sci. U. S. A. 72, 3666–3670 (1975).

26. Aggarwal, B. B. et al. Human tumor necrosis factor. Production, purification, and characterization. J. Biol. Chem. 260, 2345–2354 (1985).

27. Brenner, D., Blaser, H. & Mak, T. W. Regulation of tumour necrosis factor signalling: live or let die. Nat. Rev. Immunol. 2015 156 15, 362–374 (2015).

28. Van Loo, G. & Bertrand, M. J. M. Death by TNF: a road to inflammation. Nat Rev Immunol (2022). doi:10.1038/s41577-022-00792-3

29. Aggarwal, B. B. Signalling pathways of the TNF superfamily: a double-edged sword. Nature Reviews Immunology 745–756 (2003). Available at: https://www.nature.com/articles/nri1184.pdf. (Accessed: 29th March 2023)

30. Wang, D. & Baldwin, A. S. Activation of nuclear factor-kappaB-dependent transcription by tumor necrosis factor-alpha is mediated through phosphorylation of RelA/p65 on serine 529. J. Biol. Chem. 273, 29411–29416 (1998).

31. Farhana, A. & Khan, Y. S. Biochemistry, lipopolysaccharide. StatPearls (2022).

32. Tong, A. J., Liu, X. et al. A stringent systems approach uncovers gene-specific mechanisms regulating inflammation. Cell 165, 165–179 (2016).

33. Lu, Y. C., Yeh, W. C. & Ohashi, P. S. LPS/TLR4 signal transduction pathway. Cytokine 42, 145–151 (2008).

34. Fitzgerald, K. A. et al. LPS-TLR4 signaling to IRF-3/7 and NF-kappaB involves the toll adapters TRAM and TRIF. J. Exp. Med. 198, 1043–1055 (2003).

35. Sen, S., Cheng, Z., Sheu, K. M., Chen, Y. H. & Hoffmann, A. Gene regulatory strategies that decode the duration of NF-kappa B dynamics contribute to LPS- versus TNF-specific gene expression. Cell Syst. (2019). doi:10.1016/j.cels.2019.12.004

36. Smale, S.T. & Kadonaga, J. T. The RNA polymerase II core promoter. Annu. Rev. Biochem. 72, 449–479 (2003).

37. Andersson, R. & Sandelin, A. Determinants of enhancer and promoter activities of regulatory elements. Nat. Rev. Genet. 21, 71–87 (2020).

38. Yan, C. et al. Transcription preinitiation complex structure and dynamics provide insight into genetic diseases. 26, 397–406 (2019).

39. Wang, H., Schilbach, S., Ninov, M., Urlaub, H. & Cramer, P. Structures of transcription preinitiation complex engaged with the +1 nucleosome. Nat. Struct. Mol. Biol. | 30, 226–232 (2023).

40. Schilbach, S., Aibara, S., Dienemann, C., Grabbe, F. & Cramer, P. Structure of RNA polymerase II pre-initiation complex at 2.9 Å defines initial DNA opening. Cell 184, 4064-4072.e28 (2021).

41. Rengachari, S., Schilbach, S., Aibara, S., Dienemann, C. & Cramer, P. Structure of the human Mediator-RNA polymerase II pre-initiation complex. Nat. | 594, 129 (2021).

42. Richter, W. F., Nayak, S., Iwasa, J. & Taatjes, D. J. The Mediator complex as a master regulator of transcription by RNA polymerase II. Nat. Rev. Mol. Cell Biol. 2022 2311 23, 732–749 (2022).

43. Sen, R. & Baltimore, D. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. Cell 46, 705–716 (1986).

44. Basak, S., Shih, V. F. & Hoffmann, A. Generation and activation of multiple dimeric transcription factors within the NF-kappaB signaling system. Mol. Cell. Biol. 28, 3139–3150 (2008).

45. O'Dea, E. & Hoffmann, A. The regulatory logic of the NF-kappaB signaling system. Cold Spring Harb. Perspect. Biol. (2010). doi:10.1101/cshperspect.a000216

46. Sears, C., Olesen, J., Rubin, D., Finley, D. & Maniatis, T. NF-κB p105 processing via the ubiquitin-proteasome pathway. J. Biol. Chem. 273, 1409–1419 (1998).

47. Hayden, M., West, A. & Ghosh, S. NF-κB and the immune response. Oncogene (2006). doi:10.1038/sj.onc.1209943

48. Oeckinghaus, A. & Ghosh, S. The NF-kappaB family of transcription factors and its regulation. Cold Spring Harbor Perspectives in Biology 1, (2009).

49. Hayden, M. S. & Ghosh, S. Shared principles in NF-κB signaling. Cell 132, 344–362 (2008).

50. Kawai, T. & Akira, S. Signaling to NF-κB by Toll-like receptors. Trends Mol. Med. 13, 460–469 (2007).

51. Liu, T., Zhang, L., Joo, D. & Sun, S.-C. NF-κB signaling in inflammation. Signal Transduct. Target. Ther. 2, 17023 (2017).

52. Hayden, M. S. & Ghosh, S. Signaling to NF-kappaB. Genes Dev. 18, 2195–2224 (2004).

53. Jost, P. J. & Ruland, J. Aberrant NF-κB signaling in lymphoma: mechanisms, consequences, and therapeutic implications. Blood 109, 2700–2707 (2007).

54. Sun, S. C. The non-canonical NF-κB pathway in immunity and inflammation. Nat. Rev. Immunol. 2017 179 17, 545–558 (2017).

55. Shih, V. F., Tsui, R., Caldwell, A. & Hoffmann, A. A single NF-κB system for both canonical and non-canonical signaling. Cell Res. 21, 86 (2011).

56. Beg, A. A., Sha, W. C., Bronson, R. T., Ghosh, S. & Baltimore, D. Embryonic lethality and liver degeneration in mice lacking the RelA component of NF-kappa B. Nature 376, 167–170 (1995).

57. Sha, W. C., Liou, H. C., Tuomanen, E. I. & Baltimore, D. Targeted disruption of the p50 subunit of NF-kappa B leads to multifocal defects in immune responses. Cell 80, 321–330 (1995).

58. Snapper, C. M. et al. B cells from p50/NF-kappa B knockout mice have selective defects in proliferation, differentiation, germ-line CH transcription, and Ig class switching. J. Immunol. 156, 183–191 (1996).

59. Gerondakis, S. et al. Unravelling the complexities of the NF-kappaB signalling pathway using mouse knockout and transgenic models. Oncogene 25, 6781–6781 (2006).

60. Burkly, L. et al. Expression of relB is required for the development of thymic medulla and dendritic cells. Nature 373, 531–536 (1995).

61. Caamaño, J. H. et al. Nuclear factor (NF)-kappa B2 (p100/p52) is required for normal splenic microarchitecture and B cell-mediated immune responses. J. Exp. Med. 187, 185–196 (1998).

62. Grumont, R. J., Rourke, I. J. & Gerondakis, S. Rel-dependent induction of A1 transcription is required to protect B cells from antigen receptor ligation-induced apoptosis. Genes Dev. 13, 400 (1999).

63. Köntgen, F. et al. Mice lacking the c-rel proto-oncogene exhibit defects in lymphocyte proliferation, humoral immunity, and interleukin-2 expression. Genes Dev. 9, 1965–1977 (1995).

64. O'Dea, E. & Hoffmann, A. The regulatory logic of the NF-kappaB signaling system. Cold Spring Harb. Perspect. Biol. 2, (2010).

65. Liou, H. C., Sha, W. C., Scott, M. L. & Baltimore, D. Sequential induction of NF-κB/Rel family proteins during B-Cell terminal differentiation. Mol. Cell. Biol. 14, 5349–5359 (1994).

66. Brignall, R., Moody, A. T., Mathew, S. & Gaudet, S. Considering abundance, affinity, and binding site availability in the NF-κB target selection puzzle. Front. Immunol. 10, 609 (2019).

67. Hoffmann, A., Natoli, G. & Ghosh, G. Transcriptional regulation via the NF-κB signaling module. Oncogene 25, 6706–6716 (2006).

68. Ghosh, G., Van Duyne, G., Ghosh, S. & Sigler, P. B. Structure of NF-κB p50 homodimer bound to a κB site. Nat. 1995 3736512 373, 303–310 (1995).

69. Chen, F. E., Huang, D. Bin, Chen, Y. Q. & Ghosh, G. Crystal structure of p50/p65 heterodimer of transcription factor NF-κB bound to DNA. Nat. 1998 3916665 391, 410–413 (1998).

70. Kontgen, F. et al. Mice lacking the c-rel proto-oncogene exhibit defects in lymphocyte proliferation, humoral immunity, and interleukin-2 expression. Genes Dev. 9, 1965–1977 (1995).

71. Grinberg-Bleyer, Y. et al. NF-κB c-Rel is crucial for the regulatory T cell immune checkpoint in cancer. Cell 170, 1096-1108.e13 (2017).

72. Huang, D. B., Chen, Y. Q., Ruetsche, M., Phelps, C. B. & Ghosh, G. X-ray crystal structure of proto-oncogene product c-Rel bound to the CD28 response element of IL-2. Structure 9, 669–78 (2001).

73. Verweij, C. L., Geerts, M. & Aarden, L. A. Activation of interleukin-2 gene transcription via the T-cell surface molecule CD28 is mediated through an NF-κB-like response element. J. Biol. Chem. 266, 14179–14182 (1991).

74. Sanjabi, S. et al. A c-Rel subdomain responsible for enhanced DNA-binding affinity and selective gene activation. Genes Dev. 19, 2138–51 (2005).

75. Cao, S., Zhang, X., Edwards, J. P. & Mosser, D. M. NF-kappaB1 (p50) homodimers differentially regulate pro- and anti-inflammatory cytokines in macrophages. J. Biol. Chem. 281, 26041–26050 (2006).

76. Elsharkawy, A. M. et al. The NF-kappaB p50:p50:HDAC-1 repressor complex orchestrates transcriptional inhibition of multiple pro-inflammatory genes. J. Hepatol. 53, 519–527 (2010).

77. Yu, Y., Wan, Y. & Huang, C. The biological functions of NF-κB1 (p50) and its potential as an anti-cancer target. Curr. Cancer Drug Targets 9, 566 (2009).

78. Ghosh, G. et al. The specificity of innate immune responses is enforced by repression of interferon response elements by NF-κB p50. 11 (2011). doi:10.1126/scisignal.2001501

79. Kohda, A., Yamazaki, S. & Sumimoto, H. DNA element downstream of the κB site in the Lcn2 promoter is required for transcriptional activation by IκBζ and NF-κB p50. Genes Cells 19, 620–628 (2014).

80. Motoyama, M., Yamazaki, S., Eto-Kimura, A., Takeshige, K. & Muta, T. Positive and negative regulation of nuclear factor-kappaB-mediated transcription by IkappaB-zeta, an inducible nuclear protein. J. Biol. Chem. 280, 7444–51 (2005).

81. Trinh, D. V., Zhu, N., Farhang, G., Kim, B. J. & Huxford, T. The nuclear IκB protein IκBζ specifically binds NF-κB p50 homodimers and forms a ternary complex on κB DNA. J. Mol. Biol. 379, 122–135 (2008).

82. Yamamoto, M. et al. Regulation of Toll/IL-1-receptor mediated gene expression by the inducible nuclear protein IκBzeta. Nature 430, (2004).

83. Siggers, T. et al. Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-κB family DNA binding. Nat. Immunol. 13, 95–102 (2012).

84. Hayden, M. S. & Ghosh, S. NF-κB, the first quarter-century: remarkable progress and outstanding questions. Genes Dev. (2012). doi:10.1101/gad.183434.111

85. Baeuerle, P. A. & Baltimore, D. IκB: a specific inhibitor of the NF-κB transcription factor. Science (80-.). 242, 540–546 (1988).

86. Schuster, M., Annemann, M., Plaza-Sirvent, C. & Schmitz, I. Atypical IκB proteins-nuclear modulators of NF-κB signaling. (2013).

87. Kawai, T. & Akira, S. Signaling to NF-κB by Toll-like receptors. Trends Mol. Med. 13, 460–469 (2007).

88. Kitamura, H., Kanehira, K., Okita, K., Morimatsu, M. & Saito, M. MAIL, a novel nuclear I kappa B protein that potentiates LPS-induced IL-6 production. FEBS Lett. 485, 53–56 (2000).

89. Maruyama, T., Sayama, A., Ishii, K. J. & Muta, T. Screening of posttranscriptional regulatory molecules of IκB-ζ. Biochem. Biophys. Res. Commun. 469, 711–715 (2016).

CHAPTER 2

Impact of Chromatin Immunoprecipitation Crosslinking Conditions on Selective Genomic Interactions by NF-κB Family Members

Allison E. Daly,[1,2] Allison Schiffman,[1,2,5] Alexander Hoffmann,[1,2,5] and Stephen T. Smale[1-4*]

[1]Department of Microbiology, Immunology, and Molecular Genetics
[2]Molecular Biology Institute
[3]Department of Medicine
[4]Howard Hughes Medical Institute
[5]Institute for Quantitative and Computational Biosciences
University of California, Los Angeles, CA 90095, USA
[*]Correspondence: *smale@mednet.ucla.edu* (Tel: 310-206-4777; Fax: 310-206-8623)

SUMMARY

The ability to capture protein: DNA interactions in a quantitative, high-resolution, and temporal manner in the cell context is crucial for elucidating the mechanisms underlying transcriptional regulation. Chromatin immunoprecipitation followed by sequencing (ChIP-seq) is a widely used technique that enables the study of histone modifications and genome-wide interactions of transcription factors and chromatin regulators. Although ChIP-seq generates a vast amount of information, crosslinking conditions need to be carefully evaluated to minimize the capture of non-specific interactions. In this study, we conduct an in-depth analysis to evaluate the impact of crosslinking conditions on ChIP-seq results for two NF-κB family members, RelA and c-Rel. We titrate two commonly used chemical crosslinkers, formaldehyde and DSG. We demonstrate that while raising the crosslinking concentrations increases the total number of binding sites and the strength of the binding signal, high crosslinking conditions lead to an abundance of interactions detected at locations that lack anticipated motifs or annotations to potential NF-κB target genes. Our findings emphasize the importance of a comprehensive evaluation of ChIP-seq results and a cautious evaluation of chemical crosslinking conditions. Moreover, we compare ChIP-seq to CUT&TAG, a newer technique that does not require crosslinking. This comparison revealed strongly overlapping genomic interactions by the two methods. However, both methods display evidence of a large number of non-specific signals that exhibit little overlap between the two methods.

INTRODUCTION

Proteins interact with DNA to support biological processes such as gene transcription, DNA replication, and DNA stability. A wide range of proteins is known to interact with DNA including general transcription factors (TFs), sequence-specific TFs, histones, DNA polymerases, and RNA polymerases[1–3]. Whereas histone modifications are relatively stable, transcription factor binding to DNA is often transient and more challenging to capture. Accurately capturing the dynamic interactions of TFs with DNA is of particular interest because TFs can directly regulate transcription by binding to promoter or enhancer regions where they recruit co-activators, stabilize the pre-initiation complex, or block transcriptional activation. Because of the biological importance of understanding protein: DNA interactions, there has been considerable effort to accurately capture protein: DNA interactions in vivo at a genome-wide scale.

In 1984, chromatin-immunoprecipitation (ChIP) was pioneered in bacteria[4]. The first ChIP experiments in both prokaryotes and eukaryotes involved crosslinking via UV irradiation, followed by chromatin shearing, immunoprecipitation of RNA polymerase II, and DNA detection with a Southern blot[5,6]. Since the initial ChIP experiments, DNA detection has greatly improved from Southern blotting to quantitative PCR and sequencing. Current ChIP-seq protocols typically use chemical crosslinkers (DSG and PFA), followed by focused ultrasonication, immunoprecipitation, and next-generation sequencing[7–11]. ChIP-seq allows for the unprecedented ability to explore protein: DNA interactions on a genome-wide scale within the cell context.

In 2019, a novel method known as Cleavage Under Targets and Tagmentation (CUT&TAG) was developed to obtain high-resolution protein: DNA interactions within the cell context, while using far fewer cells than ChIP-seq[12]. In contrast to ChIP-seq, CUT&TAG does not require crosslinking or sonication. Instead, after the isolation of nuclei and primary and secondary antibody incubation, protein A/G fused to a Tn5 transposase interacts with the secondary antibody to cleave the DNA around the protein of interest. This releases DNA fragments bound by the protein of interest for library preparation and sequencing. This protocol allows for the capture of

protein: DNA interactions without chemical crosslinking and with fewer cells for the generation of high-quality datasets.

One major challenge in studying transcription factor binding at a genome-wide scale is that most transcription factors exhibit thousands of binding sites across the genome when examined by either ChIP-seq or CUT&TAG. It remains unknown whether most or all of these binding events are functionally meaningful or whether a high percentage might have no functional relevance. For example, one possibility is that many transcription factors "sample" regions of open chromatin through non-specific or low-specificity interactions. However, they may only contribute to promoter or enhancer activity at a small fraction of these regions. Although functional interactions by transcription factors are thought to often be transient, they may nevertheless require a higher binding affinity with a specific recognition sequence. Functional interactions by some sequence-specific transcription factors may occur at locations lacking a consensus binding motif, perhaps promoted by protein-protein interactions or relatively non-specific protein-DNA interactions.

In the gene regulation field, a question that remains unanswered is whether distinct members of a given TF family are capable of regulating unique sets of genes in response to identical stimuli. TFs are grouped into families based on their functional and structural properties. One TF family that is thought of as a master regulator of the immune system is the NF-κB family[13,14]. This family contains five members, c-Rel, RelA, RelB, p50, and p52, which are all expressed in immune cells and can combine to form up to 15 possible homo- and hetero-dimeric complexes[15,16]. While distinct DNA motifs for different dimers of NF-κB have been demonstrated in cell-free assays, the functional relevance of these binding preferences in vivo remains poorly understood[17–21].

Although crosslinker concentrations during ChIP-seq are often relatively high, toward the goal of capturing transient interactions, thorough investigations to understand how increasing crosslinker concentrations affect overall ChIP-seq results at a genome-wide scale have, to our knowledge, not been described in the literature. In this study, we perform in-depth analyses to

understand how increasing crosslinking affects the overall number of binding sites; the specificity of binding sites; and the functional relevance of binding sites to transcription, with a focus on NF-κB family members. We first examined the impact of ChIP-seq crosslinking conditions on the number and locations of NF-κB genomic interactions. These studies revealed that, at low concentrations of a protein-protein chemical crosslinker, thousands of genomic interactions are detected, with a strong preference for locations containing an NF-κB consensus binding motif. However, as crosslinker concentrations are increased to concentrations that are often used for the study of NF-κB, a much larger number of interactions appear, but generally at locations lacking an NF-κB consensus motif. To determine the extent to which chemical crosslinking influences the results obtained, we compared the NF-κB ChIP-seq results to results obtained with the CUT&TAG method, which does not involve crosslinking and is based on a different approach. This comparison revealed a similar number of common interaction sites, with strong enrichment for NF-κB consensus motifs, suggesting that these two distinct methods are both capable of successfully capturing specific genomic interactions. However, CUT&TAG also revealed a large number of sites that did not overlap with the ChIP-seq interaction sites, and these additional sites rarely contained NF-κB consensus motifs. These results suggest that both the crosslinking-dependent ChIP-seq method and the crosslinking-independent CUT&TAG method detect non-specific sites and, because they do not overlap between the two methods, they may represent backgrounds of different origins.

RESULTS

Protein: Protein Crosslinker DSG

To systematically investigate the impact of crosslinking on ChIP-seq results, we began by titrating chemical crosslinkers, which are used to improve the stability of protein: DNA interactions. ChIP-seq protocols for NF-κB proteins typically use two chemical crosslinkers, the protein-protein crosslinker, disuccinimidyl glutarate (DSG), and the protein-nucleic acid crosslinker, paraformaldehyde (PFA). DSG is a homobifunctional protein: protein crosslinker that covalently

binds two proteins within 10-20 Å. It is thought to improve ChIP-seq results by crosslinking the protein of interest to nearby DNA-associated chromatin proteins, thereby allowing genomic interactions by the protein of interest to be captured without the need for direct protein-DNA crosslinking by paraformaldehyde. While increasing the incubation time and temperature of DSG crosslinking influenced ChIP-seq results. We found the largest impact when titrating the concentration of chemical crosslinkers (data not shown).

To carefully examine the impact of DSG concentration on ChIP-seq results, ChIP-seq was performed with antibodies for two NF-κB family members, RelA and c-Rel, in mouse bone marrow-derived macrophages (BMDMs) stimulated with lipid A for 0 or 1 hr. Because current ChIP-seq protocols recommend the use of 0-5mM DSG, our comparison was performed with 0, 1, 2, and 4 mM DSG, in each case combined with 1% paraformaldehyde (PFA). The results revealed a dramatic increase in the number of ChIP-seq peaks with increasing DSG concentrations (Figure 2-1A). The total number of RelA peaks (p-value < 0.01) is 8,542 at 0mM DSG, which increases to 73,252 peaks at 4mM DSG. In the 1, 2, and 4mM DSG conditions the peaks are separated into two groups: 1) peaks that are called with both the condition of interest and with any lower concentrations of DSG, termed *previous peaks* and 2) peaks that are called only with the DSG concentration of interest, termed *new peaks*. Higher concentrations of DSG capture a high percentage of peaks called at lower concentrations. However, since there is not complete overlap, the number of previous peaks is not identical to the number of peaks in the lower concentration.

An example near the *Il12b* promoter further clarifies these terms (Figure 2-1D). The peak ~1,000 bps upstream of the *Il12b* TSS (highlighted in Figure 2-1D) is not detected as a statistically significant peak at 0.0mM for c-Rel or RelA. As the crosslinker concentration is increased, a peak for both c-Rel and RelA is detected and increases in score. This peak is referred to as *newly called* at the 1.0mM conditions and as *previously called* in the 2.0 and 4.0mM conditions (Figure 2-1D). We separate the peaks into these two groups to understand how peaks called at higher concentrations of DSG compare to peaks called at lower concentrations of DSG.

To determine the relationship between the peak score at newly called versus previously called peaks, we ranked the peaks for each DSG concentration based on peak score and then grouped the peaks into bins containing 1,000 peaks each. This analysis reveals that newly called peaks at higher concentrations of DSG exhibit the lowest peak scores (Figure 2-1B). For example, when investigating peaks called at 2.0mM DSG, we found that newly called peaks account for 98% of peaks in the bin with the lowest scores and 0% of the peaks in the bin with the highest scores (Figure 2-1B). This finding suggests that, while more peaks are observed with higher concentrations of DSG, these captured interactions may represent weaker genomic interactions by the RelA and c-Rel proteins. Their functional relevance may therefore require even closer scrutiny.

To further understand how the peak score at each genomic peak changes with an increasing concentration of DSG, we prepared scatter plots (Figure 2-1C). The results revealed that most peaks detected with a given DSG concentration exhibit comparable increases in peak score at the next highest concentration of DSG, with the most variability observed when the 0 and 1mM DSG concentrations were compared (i.e. a more diffuse diagonal with the 0 vs 1mM comparison than with the other comparisons). However, these scatter plots also revealed the large number of newly called peaks observed with the higher DSG concentrations that were not detected with the lower concentration (aligning with a peak score of 0 on the x-axis and with thousands of peaks overlapping at this location). While the overall correlation coefficient between neighboring conditions is high, the slope of the best-fit line reveals that, on average, the binding strength approximately doubles as the DSG condition increases (data not shown). Thus, increasing the DSG concentration increases existing peak scores relatively proportionally, while also revealing a large number of new peaks, generally with low peak scores.

NF-κB Consensus Motif Enrichment with Increasing DSG Concentrations

To further explore the characteristics of genomic interactions detected with different DSG concentrations, we examined the enrichment of canonical NF-κB motifs in peaks called at

different concentrations of DSG. For the most abundant dimer of NF-κB, p50:RelA heterodimer, the consensus motif defined biochemically and on the basis of ChIP-seq results is 5'-GGGRN(Y)YYCC-3' [22,23]. We performed Homer known motif analysis on peaks separated into four groups: 1) peaks called at 0mM DSG, 2) peaks newly called at 1mM DSG (i.e. not called at 0mM), 3) peaks newly called at 2mM DSG, and 4) peaks newly called at 4mM DSG. We found that peaks called in 0mM and newly called at 1mM DSG exhibit the strongest enrichment for canonical NF-κB motifs. However, peaks newly called at 2mM and 4mM DSG concentrations have far less enrichment for NF-κB motifs (Figure 2-2A). Peaks called at 0mM DSG, but not at the other concentrations, exhibited enrichment of POU and Stat motifs, for reasons that remain unknown.

We also performed de novo motif analysis with newly called peaks at each DSG concentration. This analysis confirmed strong enrichment of consensus NF-κB motifs among 0mM DSG called peaks and newly called peaks with 1mM DSG. However, NF-κB consensus motifs were not among the top-ranked de novo motifs when examining newly called peaks with 2mM or 4mM DSG (Figure 2-2B). These results together with the results in Figure 2-1, reveal that increasing DSG concentrations greatly increased the number of called peaks in both c-Rel and RelA ChIP-seq experiments. However, the newly detected peaks generally have low peak scores and often do not coincide with NF-κB consensus binding motifs.

It may be noteworthy that the de novo motif exhibiting the greatest enrichment among newly called peaks with 2mM and 4mM DSG resembles binding motifs for bZIP proteins. The family of bZIP proteins frequently collaborates with NF-κB dimers to coordinate inducible transcription, raising the possibility that high DSG concentrations selectively capture locations where NF-κB binds DNA in the absence of a consensus motif to coordinate inducible transcription with a nearby bZIP protein. However, the very low frequency with which these peaks annotate to inducible genes (Figure 2-2C) argues against this hypothesis. A more likely hypothesis may be that high DSG concentrations capture weak, non-functional NF-κB interactions at regions of open

chromatin throughout the genome, where bZIP motifs are among the most prevalent motifs found at regions of open chromatin.

The Proximity of RelA and c-Rel ChIP-seq Peaks to Inducible Genes

To further address the characteristics of RelA and c-Rel ChIP-seq peaks observed at different DSG concentrations, we first divided newly called peaks at each DSG concentration into ten bins based on peak score. We then annotated each called peak to the nearest gene and calculated the percentage of peaks within each bin that annotated to a gene whose transcript levels increased >5-fold following stimulation with lipid A (Figure 2-2C). Induction of transcript levels was determined using previously described nascent transcript RNA-seq data from BMDMs stimulated lipid A for 0 or 1hr[24].

This analysis revealed that, with 0mM DSG, a higher percentage of peaks annotated to strongly induced genes than was observed with higher DSG concentrations (Figure 2-2C). Approximately 20% of peaks in the bin containing the highest peak scores annotated to induced genes, with much lower percentages annotating to induced genes in 0mM bins containing smaller peak scores, and in all bins with newly detected peaks obtained with 1, 2, or 4mM DSG. Notably, in the 0mM DSG bins with the smallest peak scores, 7-8% of peaks annotated near induced genes, which is larger than the 2-3% observed with peaks newly called with 4mM DSG. The 2-3% observed with 4mM DSG may represent background. Thus, if proximity to an inducible gene is viewed as a preliminary estimate of functional significance, peaks called with 0mM DSG have a high probability of functional relevance regardless of peak score compared to peaks newly called with 4mM DSG. However, even with 0mM DSG, the probability of functional significance appears to increase with increasing peak score. Notably, some peaks newly called with 1mM DSG and to a lesser extent with 2mM DSG also have a higher probability of functional significance than the apparent background, but primarily in bins with the strongest peak scores. Together, the results reveal that the large number of peaks newly called with increasing DSG concentrations, especially 2 and 4mM DSG, possesses a lower probability of functional relevance. These peaks not only are

generally weak and unlikely to coincide with consensus NF-κB motifs, but they also are far less likely to annotate to inducible genes than peaks observed with 0mM and 1mM DSG, especially peaks with large peak scores captured with 0 and 1mM DSG.

Protein: DNA Crosslinker PFA

PFA can catalyze both protein-protein and protein-DNA crosslinking but has emerged as a preferred crosslinking agent for ChIP-seq experiments due to the latter activity. PFA can easily cross cell membranes and most commonly catalyzes the formation of covalent crosslinks between lysine residues in DNA-bound proteins and DNA bases. Since standard ChIP-seq protocols include 0.5% - 2.0% PFA, we performed ChIP-seq experiments with 0.0%, 0.5%, 1.0%, and 2.0% PFA in combination with 1mM DSG for both c-Rel and RelA ChIP-seq in BMDMs stimulated with lipid A for 0 or 1hr. We first assessed the overall number of statistically called peaks (p-value < 0.01) in each condition (Figure 2-3A). In contrast to the large number of peaks obtained in the absence of DSG, only 289 and 192 peaks were observed for c-Rel and RelA, respectively, in the absence of PFA. With the addition of 0.5% PFA, there is a dramatic increase in the number of peaks, which continued to increase with higher PFA concentrations.

Similar to the DSG analyses, increasing the PFA concentration yielded many of the same peaks observed with the lower PFA concentrations, but also a large number of newly called peaks (Figure 2-3A). Also similar to the DSG results, the newly called peaks with higher PFA concentrations were generally weak (Figure 2-3B). When using scatter plots to examine the impact of increased PFA concentrations on peak scores with individual peaks, peaks observed with the lower PFA concentration generally increased their peak score with the higher concentration, in addition to the appearance of the newly called peaks (Figure 2-3C). This increase in the peak score for pre-existing peaks was most pronounced when the PFA concentration was increased from 0.5% to 1%. The *Ccl5* promoter contains an example of a peak that first appears with 0.5% PFA and continues to increase in peak score at 1% and 2% PFA (Figure 2-3D).

We also examined motif enrichment at newly called peaks with each PFA concentration, analogous to the above analysis with DSG. The most significant enrichment of a consensus NF-κB motif was observed with newly called peaks with 0.5% PFA (Figure 2-4A). This motif was also enriched in newly called peaks with 1% and 2% PFA but to a lesser extent. In a de novo motif analysis, a motif resembling an NF-κB consensus exhibited the greatest enrichment with the newly called peaks observed with 0.5% PFA (Figure 2-4B). However, motifs resembling an NF-κB consensus were much more poorly enriched among newly called peaks observed with 1% and 2% PFA, with the most enriched motif suggestive of bZIP family protein binding. Given the very small number of peaks called with 0% PFA, it is unsurprising that the most enriched motif had little resemblance to an NF-κB motif.

Finally, newly called peaks with each PFA concentration were merged with nascent transcript RNA-seq data to determine the prevalence with which the newly called peaks annotated to inducible genes. This analysis revealed that the newly called peaks observed with 0.5% PFA exhibited a relatively high probability of annotating to induced genes (Figure 2-4C). Peaks in bins with the highest scores exhibited the highest probability of annotating to inducible genes, but all peak-score bins observed with 0.5% PFA had a higher probability than any of the bins derived from 1% or 2% PFA. Together, these results suggest that, for both RelA and c-Rel, 0.5% PFA appears to capture a large fraction of ChIP-seq peaks that have the highest probability of functional relevance, as measured by annotation to induced genes (Figure 2-4C) and by NF-κB motif enrichment (Figure 2-4B) and peak score. Notably, however, one major benefit of 1% PFA is a large increase in peak score for those peaks initially detected with 0.5% PFA (see Figure 2-3C).

ChIP-seq Versus CUT & TAG Comparison

Given the large effect crosslinking conditions have on the outcome of a ChIP-seq experiment, we next explored how ChIP-seq compares to CUT&TAG, an assay that does not use crosslinking. Instead of crosslinking and sonication, CUT&TAG employs a protein A/G-Tn5 transposase fusion

protein that tethers to a secondary antibody and can cleave the DNA around a protein of interest. This releases DNA that can then be prepared for sequencing. CUT&TAG has been well described for the capture of histone modifications and is being used with increasing frequency for the analysis of transcription factor interactions. We, therefore, compared ChIP-seq and CUT&TAG for the NF-κB member, RelA. We began by comparing the number of overlapping peaks called with each technique, including all of the crosslinker concentrations for ChIP-seq. The percentage shown in the heatmap corresponds to the number of overlapping peaks versus the number of peaks at each condition (Figure 2-5A; peak numbers at the top).

Focusing first on the different ChIP-seq crosslinker concentrations, the heatmap recapitulates the results shown above, in that >90% of peaks called at a low crosslinker concentration overlap with peaks called at higher crosslinking conditions. For example, of the 2,665 RelA peaks observed with 0mM DSG, almost all are observed with 1mM, 2mM, and 4mM DGS (Figure 2-5A, top left). In contrast, of the 15,819 peaks observed with 2mM DSG, relatively small percentages are observed with 0mM or 1mM DSG, but almost all are observed with 4mM DSG (Figure 2-5A, top left).

Next, a comparison between the ChIP-seq and CUT&TAG results reveals substantial overlap. For example, at all three CUT&TAG time points, a substantial percentage of CUT&TAG peaks overlap with the ChIP-seq peaks observed with all four DSG concentrations (Figure 2-5A, top right). Although almost all CUT&TAG peaks are contained within the 64,824 peaks obtained with 4mM DSG, this is likely to be because these 64,824 peaks represent a high percentage of regions of open chromatin throughout the genome.

Using discrete cutoffs to examine more closely the overlap between ChIP-seq DSG titration data and CUT&TAG data, we found substantial overlap, but we also observed thousands of called peaks that are unique to either ChIP-seq or CUT&TAG (Figure 2-5B, top). For example, in a comparison of 1mM DSG ChIP-seq data with the CUT&TAG data, 3,638 peaks overlap, representing 38% and 55% of the ChIP-seq and CUT&TAG peaks, respectively (Figure 2-5B, top). This extensive overlap using two entirely different methodologies suggests that both techniques

successfully capture bona fide NF-κB genomic binding sites. However, 62% and 45% of the ChIP-seq and CUT&TAG peaks, respectively, do not overlap with peaks obtained with the other technique.

To gain insight into the characteristics of the peaks detected by the two methods, we performed motif analysis with peaks detected by both methods and peaks detected exclusively with one method (Figure 2-5B, bottom). Peaks detected in common by both methods showed the strongest enrichment of NF-κB motifs, regardless of whether the 0mM, 1mM, or 2mM DSG data was used for the analysis, suggesting that interactions detected with these two distinct methods have the highest probability of representing bona fide NF-κB interaction sites (Figure 2-5B, bottom). In contrast, weaker enrichment of NF-κB motifs was found with peaks observed in only one of the two techniques, with the weakest enrichment of NF-κB motifs observed at the collection of peaks detected using only the CUT&TAG method. These results suggest that both methods are susceptible to the capture of interactions that are less likely to reflect the specific binding of NF-κB to its consensus recognition motifs.

DISCUSSION

In this study, we examined at a genome-wide scale how NF-κB protein: DNA interactions detected by ChIP-seq are impacted by crosslinker concentrations and how NF-κB ChIP-seq and CUT&TAG results compare with each other. Although crosslinker titrations are frequently performed when developing a ChIP-seq assay for a new protein, these titrations are generally evaluated by PCR with only a small number of representative binding sites. We are unaware of other published reports describing the impact of crosslinker titrations at a genome-wide scale. We titrated two chemical crosslinkers commonly used in ChIP-seq and evaluated the impact on the overall ChIP-seq results and the specificity of interactions. The comparison between ChIP-seq and CUT&TAG results provided an opportunity to not only compare two entirely different methodologies but to compare a crosslinking-dependent method with a method that does not

involve crosslinking. The results of these analyses reveal interesting insights and underscore the value of thorough analyses when studying protein-DNA interactions.

Given that the NF-κB family is well studied, has a defined consensus motif, and is a prominent regulator of inducible transcription in immune cells, we used two members of the NF-κB family, RelA and c-Rel, as models for these studies. Because c-Rel and RelA homodimers bind similar ranges of recognition motifs but with highly divergent binding affinities[25], we originally thought the impact of crosslinker concentration might be different with the two proteins. However, when titrating crosslinker concentrations, we saw highly similar changes for both c-Rel and RelA.

The results described here demonstrate that two chemical crosslinkers commonly used in ChIP-seq (DSG and PFA) have dramatic effects on the total number of binding sites. Toward the goal of understanding if additional peaks called only at higher concentrations are functional, we explored the motif enrichment for NF-κB motifs and the enrichment of peaks that annotate to inducible genes. This approach allowed us to decipher whether peaks called only at higher concentrations of crosslinker have characteristics suggestive of functional interactions. We found that, at high conditions of DSG (2mM, 4mM) and PFA (2.0%), there is a lack of enrichment of consensus NF-κB motifs. Furthermore, peaks called exclusively at higher crosslinker concentrations have a low probability of annotating to inducible genes. These results suggest that increasing the crosslinker concentration often leads to the capture of “off-target” or background interactions. Therefore, for any given factor that interacts with DNA, there may be defined crosslinker concentrations that allow for the optimal capture of meaningful protein-DNA interactions, while minimizing to the greatest extent possible the capture of aberrant interactions. Importantly, further analyses are needed to understand the extent to which interactions captured by “optimal” crosslinker concentrations are functionally relevant, despite the strong enrichment of consensus motifs when using these conditions. Further analyses are also needed to determine

whether functional relevance can be attributed to the additional interactions captured with only higher crosslinker concentrations.

The substantial overlap in interactions captured by ChIP-seq and CUT&TAG reinforces the previous evidence that CUT&TAG is a highly valuable technique, given the method's ease of use and suitability for use with small numbers of cells. However, both ChIP-seq and CUT&TAG captured large numbers of interactions that were not detected with the other method, and NF-κB consensus motifs exhibited much weaker enrichment at peaks captured with only one of the two methods. Although the significance of this finding is not known, a likely possibility is that each method has distinct susceptibilities to capturing interactions that are less likely to represent specific protein-DNA interactions or to be functionally important, in comparison to peaks that are reproducibly captured by both methods.

The use of chemical crosslinkers for ChIP-seq has long been known to provide an opportunity for background, due to the ability of a crosslinker to fix a highly transient interaction that may occur with little or no specificity. This concern is eliminated with CUT&TAG. However, the CUT&TAG method is susceptible to different potential challenges, including the potential for the transposase to preferentially cleave DNA at genomic sites without prior binding to the antibody bound to the protein of interest. Optimization of experimental conditions for CUT&TAG will help minimize these potential background cleavage events, but effective approaches will be needed to evaluate the results of optimization experiments. As done for the current analysis, evaluation of the optimal enrichment of consensus recognition motifs for a transcription factor of interest, or optimal enrichment of interactions near potential target genes, may be the preferred approach. However, this strategy will be of limited value for those transcription factors that may frequently carry out functional interactions with sites that diverge from their in vitro consensus recognition sequence, or for factors that may regulate target genes that diverge from those that can be predicted on the basis of current knowledge.

EXPERIMENTAL PROCEDURES

Cell Culture

Bone marrow-derived macrophages (BMDMs) were prepared from C57BL/6 male mice 8-12 weeks of age. Following extraction of the bone marrow, cells were incubated for 4 days with 10% CMG to begin differentiation into macrophages. On day 4, cells were scraped and plated at 5x10^6 cells per 15cm plate in fresh media containing 10% CMG. On day 6, cells were stimulated with 100ng/mL lipid A (Sigma) for 60 mins.

RNA-seq

Chromatin-associated RNA-seq data used in these studies were prepared as previously described (Bhatt et al., 2012). 100ng of chromatin-associated RNA-seq was used for library preparation with the TruSeq stranded RNA kit. Libraries were sequenced using HiSeq 3000 for a single-end read length of 50bps. The read depth was approximately 30 million reads per sample. Following demultiplexing, reads were aligned with Hisat2 to the NCBI37/mm9 genome. RPKM values were calculated for each sample by dividing the total reads for one gene by the length of the gene in kbps and the total reads per sample. Gene induction was calculated by averaging the RPKM values for three replicates at the 0.0h and 1.0h LPS stimulation time points and taking the ratio of 1.0h avg. RPKM versus the 0.0h avg RPKM.

ChIP-seq

ChIP-seq was performed as previously described (Barish et al., 2010; Lee et al., 2006) with anti-RelA antibody (Cell Signaling, 8242), anti-c-Rel antibody (Cell Signaling, 67489), or anti-p50 antibody (Cell Signaling, 13586). Approximately 10 million BMDMs were used per sample. After cross-linking with DSG (0mM – 4mM) and PFA (0% - 2.0%), cells were sonicated on a Covaris

M220-focused ultrasonicator. The proper distribution of chromatin was checked using DNA electrophoresis to ensure DNA fragmentation was between 200-500bps.

ChIP-seq libraries were prepared using KAPA HyperPrep Kits (Roche) and barcode indices from NextFlex (Perkin Elmer). Following sequencing and demultiplexing, reads were aligned using Hisat2 to the NCIB/mm9 mouse genome. Peak calling was performed with Homer software, using input samples to find peak enrichment with a p-value. < 0.01 (Heinz *et al.,* 2010). To compare peaks across multiple samples, a master probe was generated with BEDTools (Quinlan et al., 2010). Then RPKMs were generated using raw reads from SeqMonk normalized to the size of the peak (in kbps) and the depth of sample sequencing (in million reads).

CUT&TAG

CUT&TAG was performed with an anti-RelA antibody (Santa Cruz sc-109) as previously described (Kaya-Okur et al., 2019, & EpiCypher). Approximately 150,000 BMDMs were used per sample. Universal i5 and barcoded i7 primers were added to samples and library preparation was completed with NEB Next 2X PCR amplification, and Ampure Bead cleanup of samples. Sequencing was performed using HiSeq 3000, to generate 50bp single-end reads with about 10 million reads per sample. Samples were processed, aligned, and normalized with the same parameters as described with ChIP-seq.

FIGURE LEGENDS

Figure 2-1. Broad Effects of DSG on ChIP-seq

BMDMs stimulated with lipid A for 1hr are crosslinked with 0, 1, 2, or 4mM DSG during ChIP-seq. The results shown are from a single replicate of the DSG titration.

(A) The total number of peaks (Peak Score > 0) for each condition is plotted. Within each DSG condition, peaks are further grouped into two categories: peaks called only at the DSG condition

of interest (darker color) or peaks called with both the condition of interest and any lower concentration of DSG (lighter color). Analyses for RelA and c-Rel peaks are colored in orange and blue, respectively.

(B) All peaks are ranked based on peak score strength and separated into bins with 1,000 peaks per bin. The percentage of peaks that are either newly called in the concentration of interest (darker color) or called in the concentration of interest and a lower concentration (lighter color) are plotted.

(C) The scatter plots show the peak score at each concentration of DSG. The plots show all peaks that are bound in either condition with a peak score > 0.

(D) The IGV genome browser track is shown for c-Rel and RelA ChIP-seq with 0, 1, 2, or 4mM DSG. A peak that is not called at the 0mM concentration for c-Rel and RelA but is called at higher concentrations is highlighted in yellow.

Figure 2-2. Assessing the Functional Relevance of Titrating DSG

(A) All peaks are categorized into four groups. 1) Peaks called at 0mM DSG (Peak Score at 0mM > 0); 2) newly called peaks at 1mM (Peak Score at 1mM > 0 & Peak Score at 0mM = 0); 3) newly called peaks at 2mM (Peak Score at 2mM > 0 & Peak Score at 1mM = 0), and 4) newly called peaks at 4mM (Peak Score at 4mM > 0 & Peak Score at 2mM = 0). A random sample of 1,000 peaks was used for Homer known motif analysis. The heat map corresponds to the –log10(p-value).

(B) All peaks in the four categories previously described were analyzed with Homer De Novo motif analysis software. The top motif in each analysis is shown.

(C) All peaks are divided into the four categories as previously described. Every peak is annotated to the nearest genes. Peaks in each category are ranked based on peak score and grouped into 10 bins. The percent of peaks in each bin that annotate to an inducible gene (Induction > 5) are plotted.

Figure 2-3. Broad Effects of PFA on ChIP-seq

BMDMs stimulated with lipid A for 1hr are crosslinked with 0, 0.5, 1.0, or 2.0% PFA during ChIP-seq. The results shown are from a single replicate of the PFA titration.

(A) The total number of peaks (Peak Score > 0) for each condition is plotted. Within each PFA condition, peaks are further grouped into two categories: peaks called only at the PFA condition of interest (darker color) or peaks called with both the condition of interest and any lower concentration of PFA (lighter color). Analyses for RelA and c-Rel peaks are colored in orange and blue, respectively.

(B) All peaks are ranked based on peak score strength and then separated into bins with 1,000 peaks per bin. The percentage of peaks that are either newly called in the concentration of interest (darker color) or called in the concentration of interest and a lower concentration (lighter color) are plotted.

(C) The scatter plots show the peak score at each concentration of PFA. The plots show all peaks that are bound in either condition with a peak score > 0.

(D) The IGV genome browser track is shown for c-Rel and RelA ChIP-seq with 0, 0.5, 1.0, or 2% PFA. A peak that is not called at the 0% concentration for c-Rel and RelA but is called at higher concentrations is highlighted in yellow.

Figure 2-4. Assessing the Functional Relevance of Titrating PFA

(A) Peaks are categorized into four groups. 1) Peaks called at 0% PFA (Peak Score 0% > 0); 2) newly called peaks at 0.5% PFA (Peak Score at 0.5% > 0 & Peak Score at 0% = 0); 3) newly called peaks at 1.0% (Peak Score at 1.0% > 0 & Peak Score at 0.5% = 0), and 4) newly called peaks at 2.0% (Peak Score at 2.0% > 0 & Peak Score at 1.0% = 0). A random sample of 1,000 peaks was used for Homer known motif analysis. The heat map corresponds to the –log10(p-value).

(B) All peaks in the four categories previously described were analyzed with Homer De Novo motif analysis software. The top motif in each analysis is shown.

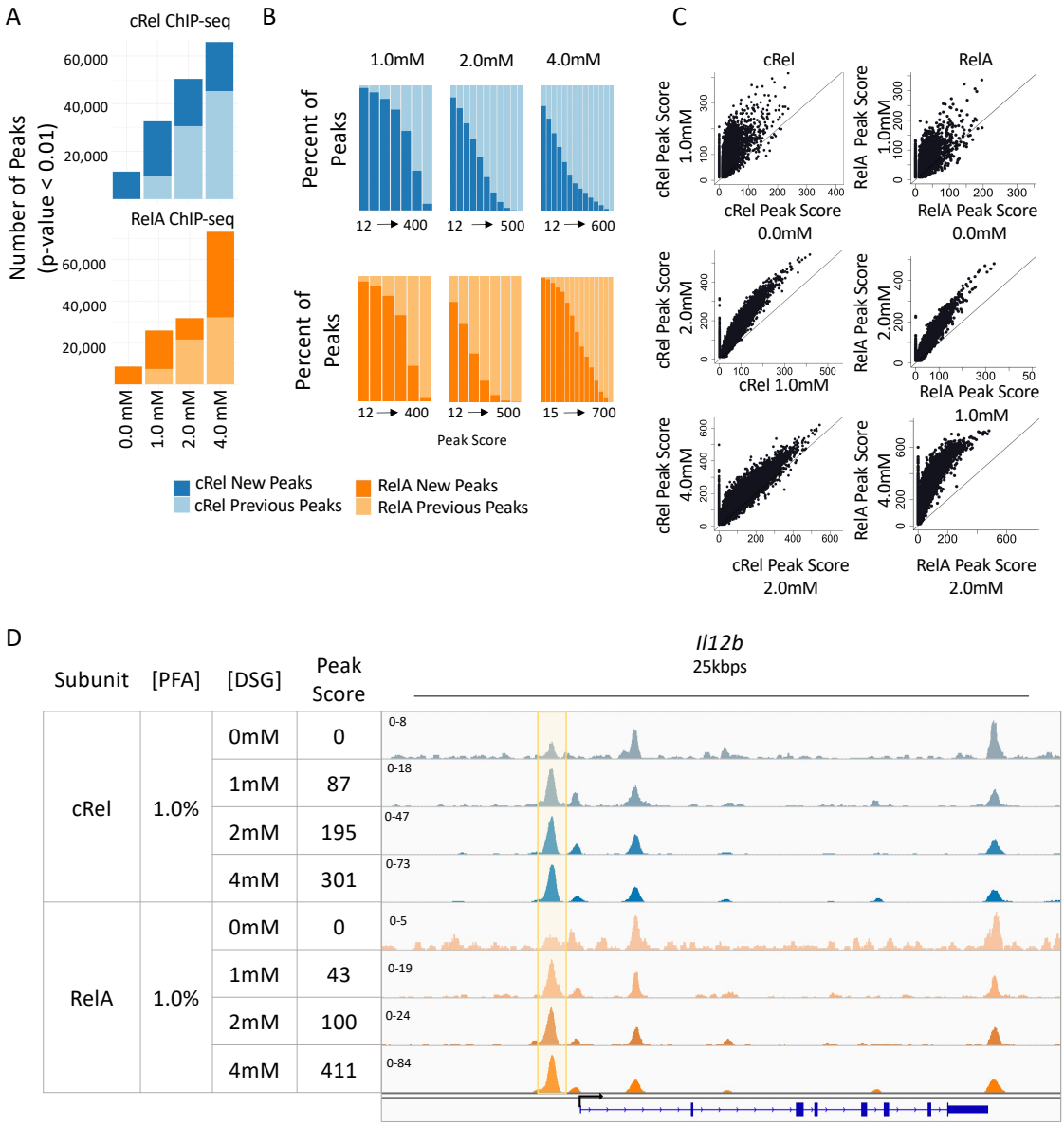
(C) All peaks are divided into the four categories as previously described. Every peak is annotated to the nearest genes. Peaks in each category are ranked based on peak score and further grouped into 10 bins. The percent of peaks in each bin that annotate to an inducible gene (Induction > 5) are plotted.

Figure 2-5: ChIP-seq versus CUT&TAG Comparison

(A) The heat map shows the percent overlap between each ChIP-seq sample and CUT&TAG samples. The number of peaks that overlap between the two assays relative to the number of peaks in each sample (listed at the top of the chart) are shown as percentages in the heatmap.

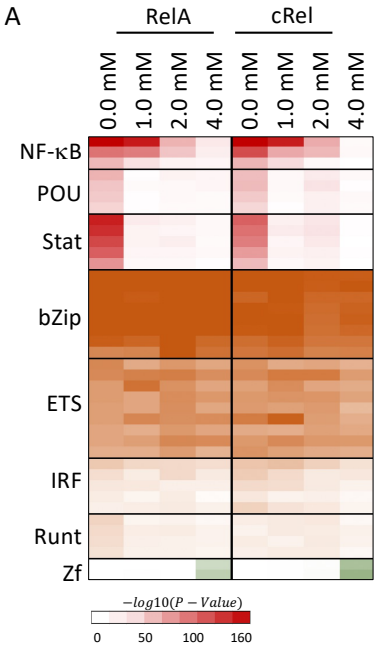
(B) The overlap between the ChIP-seq DSG titrations and the CUT&TAG datasets are shown as Venn diagrams (top). Known Homer motif analysis was performed on peaks grouped into three categories: 1) peaks called in ChIP-seq (p-value < 0.01) and CUT&TAG (p-value < 0.01), 2) peaks only called in ChIP-seq, or 3) peaks only called in CUT&TAG.

Figure 2-1. Broad Effects of DSG on ChIP-seq



Subunit	[PFA]	[DSG]	Peak Score
cRel	1.0%	0mM	0
		1mM	87
		2mM	195
		4mM	301
RelA	1.0%	0mM	0
		1mM	43
		2mM	100
		4mM	411

Figure 2-2. Assessing the Functional Relevance of Titrating DSG



B

[DSG]	cRel ChIP-seq:	% of peaks with motifs
0.0 mM		32%
1.0 mM		31%
2.0 mM		28%
4.0 mM		28%

[DSG]	RelA ChIP-seq:	% of peaks with motifs
0.0 mM		34%
1.0 mM		30%
2.0 mM		26%
4.0 mM		21%

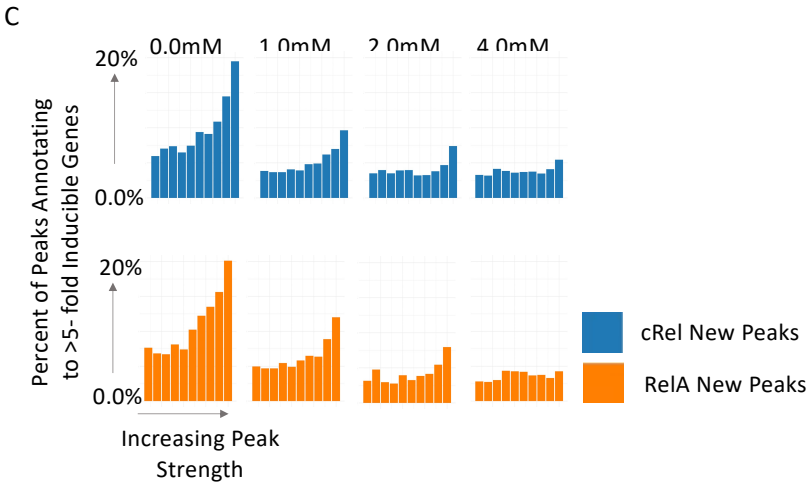
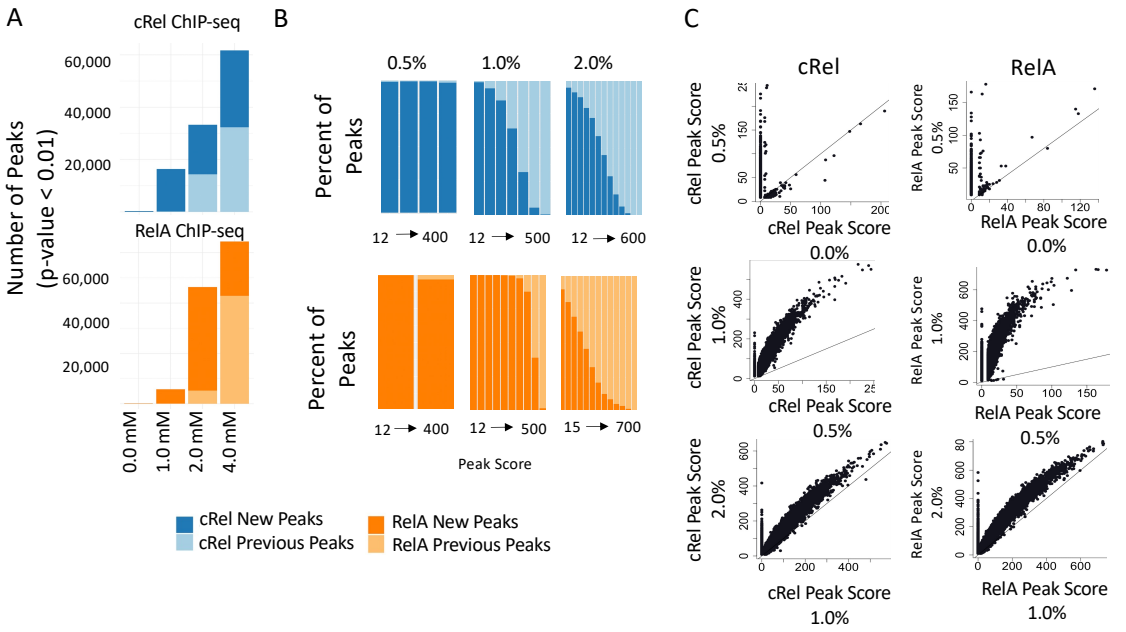


Figure 2-3. Broad Effects of PFA on ChIP-seq



D

Subunit	[DSG]	[PFA]	Peak Score
cRel	1mM	0.0%	0
		0.5%	23
		1.0%	286
		2.0%	478
RelA	1mM	0.0%	0
		0.5%	33
		1.0%	201
		2.0%	332

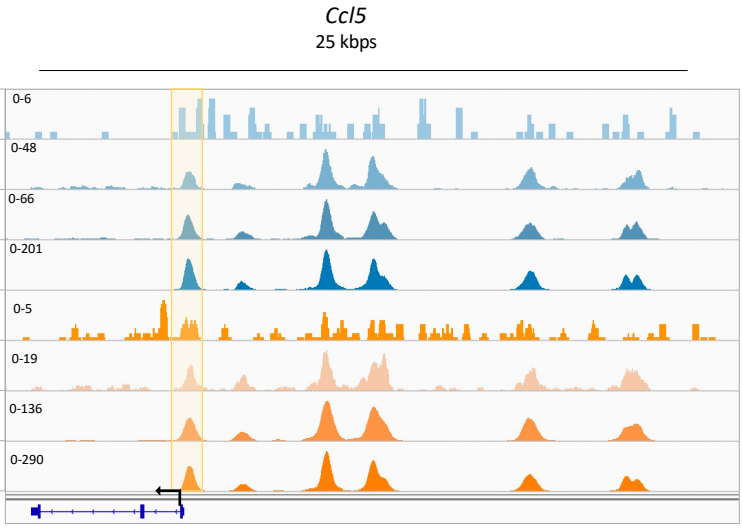


Figure 2-4. Assessing the Functional Relevance of Titrating PFA

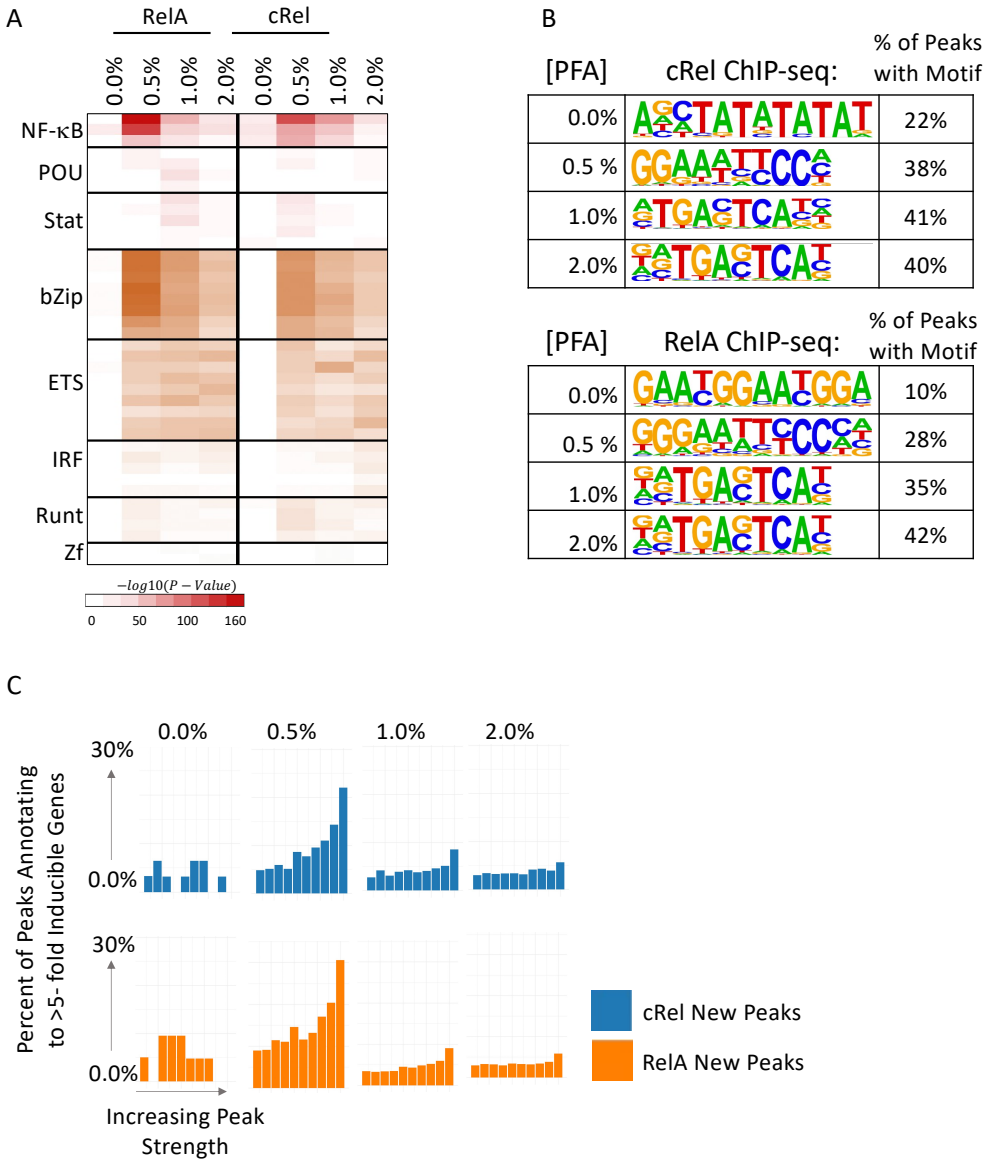
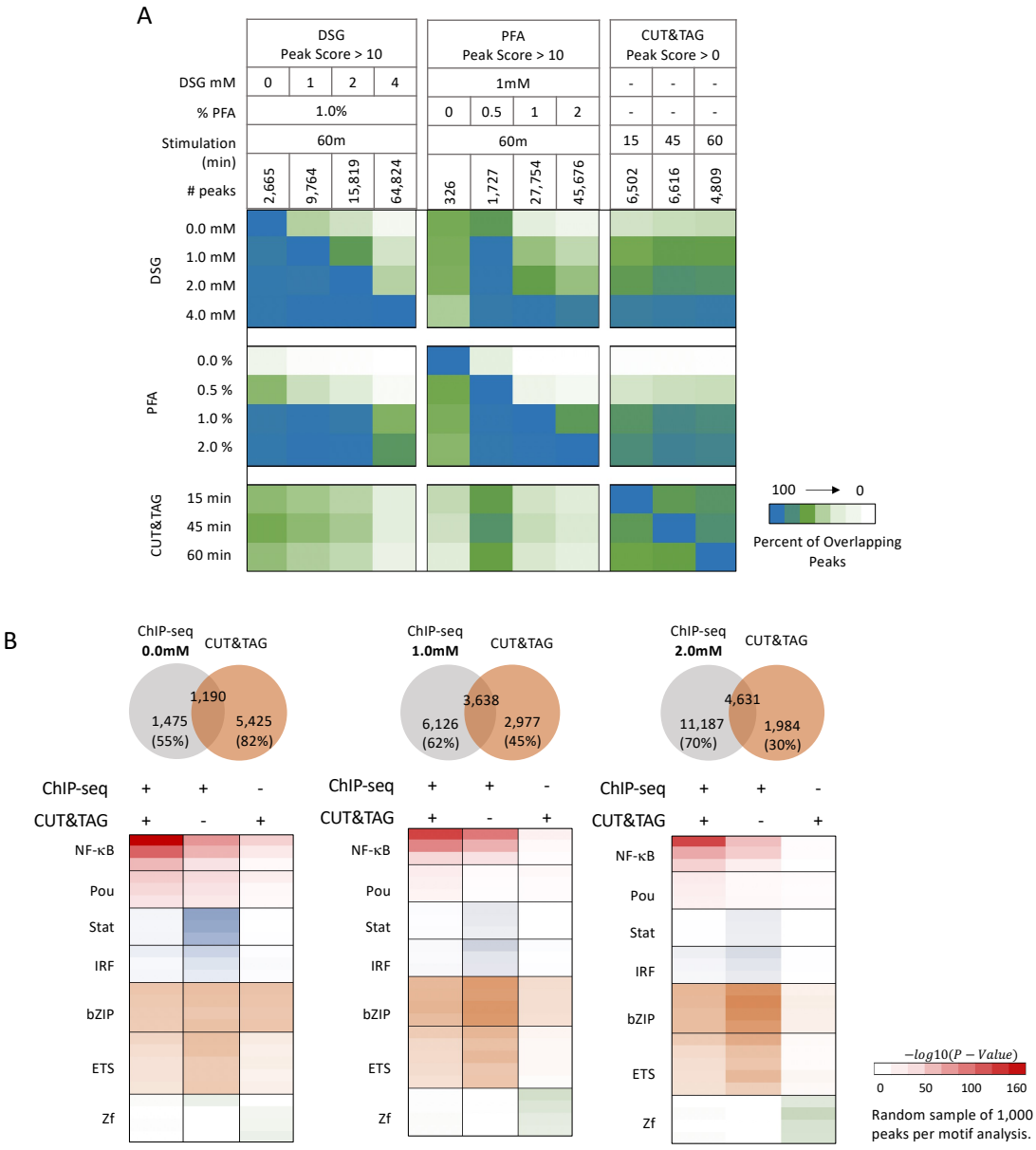


Figure 2-5. ChIP-seq versus CUT&TAG Comparison



WORKS CITED

1. Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. Cell 152, 1237–1251 (2013).
2. Stormo, G. D. & Zhao, Y. Determining the specificity of protein–DNA interactions. (2010). doi:10.1038/nrg2845
3. Campos, E. I. & Reinberg, D. Histones: annotating chromatin. https://doi.org/10.1146/annurev.genet.032608.103928 43, 559–599 (2009).
4. Gilmour, D. S. & Lis, J. T. Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. Proc. Natl. Acad. Sci. U. S. A. 81, 4275 (1984).
5. Gilmour, D. S. & Lis, J. T. RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. Mol. Cell. Biol. 6, 3984–3989 (1986).
6. Gilmour, D. S. & Lis, J. T. In vivo interactions of RNA polymerase II with genes of Drosophila melanogaster. Mol. Cell. Biol. 5, 2009 (1985).
7. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-wide mapping of in vivo protein-DNA interactions. Science (80-.). 316, 1497–1502 (2007).
8. Hoon Kim, T. & Ren, B. Genome-wide analysis of protein-DNA interactions. (2006). doi:10.1146/annurev.genom.7.080505.115634
9. Tian, B., Yang, J. & Brasier, A. R. Two-step crosslinking for analysis of protein-chromatin interactions. doi:10.1007/978-1-61779-376-9_7
10. Gade, P. & Kalvakolanu, D. V. Chromatin immunoprecipitation assay as a tool for analyzing transcription factor activity. Methods Mol. Biol. 809, 85–104 (2012).
11. de Jonge, W. J., Brok, M., Kemmeren, P. & Holstege, F. An extensively optimized chromatin immunoprecipitation protocol for quantitatively comparable and robust results. bioRxiv 835926 (2019). doi:10.1101/835926

12. Kaya-Okur, H.S. et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. doi:10.1038/s41467-019-09982-5

13. Ghosh, S. & Hayden, M. S. Celebrating 25 years of NF-κB research. Immunol. Rev. 246, 5–13 (2012).

14. Hayden, M. S. & Ghosh, S. NF-κB, the first quarter-century: remarkable progress and outstanding questions. Genes Dev. 26, 203–234 (2012).

15. O'Dea, E. & Hoffmann, A. The regulatory logic of the NF-kappaB signaling system. Cold Spring Harb. Perspect. Biol. (2010). doi:10.1101/cshperspect.a000216

16. Hoffmann, A., Natoli, G. & Ghosh, G. Transcriptional regulation via the NF-κB signaling module. Oncogene 25, 6706–6716 (2006).

17. Sanjabi, S. et al. A c-Rel subdomain responsible for enhanced DNA-binding affinity and selective gene activation. Genes Dev. 19, 2138–51 (2005).

18. Chen, F. E., Huang, D. Bin, Chen, Y. Q. & Ghosh, G. Crystal structure of p50/p65 heterodimer of transcription factor NF-κB bound to DNA. Nat. 1998 3916665 391, 410–413 (1998).

19. Ghosh, G., Van Duyne, G., Ghosh, S. & Sigler, P. B. Structure of NF-κB p50 homodimer bound to a κB site. Nat. 1995 3736512 373, 303–310 (1995).

20. Sanjabi, S., Hoffmann, A., Liou, H.-C., Baltimore, D. & Smale, S. T. Selective requirement for c-Rel during IL-12 P40 gene induction in macrophages. Proc. Natl. Acad. Sci. 97, 12705–12710 (2000).

21. Siggers, T. et al. Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-κB family DNA binding. Nat. Immunol. Vol. 13, (2012).

22. Oeckinghaus, A. & Ghosh, S. The NF-kappaB family of transcription factors and its regulation. Cold Spring Harbor Perspectives in Biology 1, (2009).

23. Hoffmann Alexander, Leung H. Thomas & Baltimore David. Genetic analysis of NF-κB/Rel transcription factors defines functional specificities. EMBO J. (2003).

24. Tong, A.-J. et al. A Stringent Systems Approach Uncovers Gene-Specific Mechanisms Regulating Inflammation. Cell 165, 165–179 (2016).

25. Siggers, T. et al. Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-κB family DNA binding. Nat. Immunol. 13, 95–102 (2012).

CHAPTER 3

Contributions of an NF-κB p50-IκBζ Pathway to the Selective Regulation of Pro-Inflammatory Gene Transcription

Allison Daly,[1,2] George Yeh,[1,2] Sofia Soltero,[1,3] and Stephen T. Smale[1-4*]

[1]Department of Microbiology, Immunology, and Molecular Genetics
[2]Molecular Biology Institute
[3]Howard Hughes Medical Institute, University of California, Los Angeles, Los Angeles, CA 90095, USA
[4]Lead contact
[*]Correspondence: smale@mednet.ucla.edu

SUMMARY

The five NF-κB family members and three nuclear IκB proteins play diverse biological roles. However, the mechanisms by which distinct NF-κB dimers and NF-κB:IκB complexes contribute to selective gene transcription at a genome-wide scale remain poorly understood. Despite multiple potential regulatory mechanisms attributed to the NF-κB p50 protein, we observed extensive overlap between p50-dependent and IκBζ-dependent pro-inflammatory genes in Toll-like receptor 4 (TLR4)-activated mouse macrophages, demonstrating more frequent collaboration than anticipated. However, p50-IκBζ co-dependence did not coincide with genomic sites that preferentially bind p50 homodimers. Instead, p50, IκBζ, and RelA co-occupied thousands of sites. A defining characteristic of sites near p50-IκBζ co-dependent genes was the binding of both IκBζ and RelA in a p50-dependent manner. In vitro and in vivo interactions between IκBζ, p50, and RelA provided further evidence that IκBζ function is not restricted to p50 homodimers. Notably, IκBζ and the IκBζ target genes comprised a high percentage of genes that exhibited the greatest differential expression between TLR4-stimulated and tumor necrosis factor receptor (TNFR)-stimulated macrophages, with transcription of a subset of these genes rescued by ectopic expression of IκBζ. These results reveal a dedicated p50-IκBζ pathway involving a previously unappreciated mechanism that selectively activates a set of key immunoregulatory genes and serves as an important contributor to the differential responses to TNFR and TLR4.

Keywords: macrophages, inflammation, transcription, NF-κB, IκBζ

INTRODUCTION

Inflammatory responses are activated in diverse mammalian cell types by a wide range of physiological stimuli, including microorganisms and microbial products, cytokines, and other environmental insults[1]. The transcriptional activation of pro-inflammatory genes, which represents a critical component of an inflammatory response, is tailored to the stimulus to defend the host against the specific insult and restore cellular and organismal homeostasis[2,3]. The tailored transcriptional response is dictated by the sensors of the stimulus; the signaling pathways and transcription factors induced by the sensors; and the poised and stimulus-responsive chromatin state of the genome[4–7].

The differential responses to the microbial product, lipid A (the active subunit of lipopolysaccharide), and the cytokine, tumor necrosis factor (TNF), have long served as a model for understanding the selectivity of pro-inflammatory gene induction. These stimuli, which are sensed by Toll-like receptor 4 (TLR4) and the TNF receptors (TNFRs), respectively, are potent inducers of common transcription factors, including NF-κB, AP-1, and serum response factor (SRF)[8–11]. However, despite the extensive overlap in transcriptional programs, the responses to TLR4 and TNFRs exhibit important differences. Notably, the type 1 interferon (IFN) response and several key cytokines and effector molecules are selectively activated by TLR4 signaling. A well-characterized contributor to this differential response is TLR4's ability to activate the TRIF signaling pathway and its downstream IRF3 transcription factor, which is critical for the type 1 IFN response[12]. The TRIF pathway also prolongs the activation of NF-κB complexes, allowing activation kinetics to serve as a major regulator of the selective transcriptional response[11].

Most studies of NF-κB's role in inflammatory gene transcription have focused on the abundant heterodimer composed of the NF-κB RelA and p50 subunits. However, the NF-κB family consists of five members – RelA, c-Rel, RelB, p50, and p52 – each of which contains a conserved Rel homology region (RHR) that supports sequence-specific DNA binding and dimerization[13–15]. The five subunits assemble into 15 homodimeric and heterodimeric species, with the RelA, c-Rel, and RelB subunits containing transactivation domains[16]. Many NF-κB dimers

are induced by post-translational mechanisms in response to inflammatory stimuli[17]. Phenotypic studies of mice lacking the genes encoding individual NF-κB subunits have revealed distinct immunological defects. However, much remains to be learned about dimer-specific functions in vivo and the underlying mechanisms of dimer-specific regulation.

The p50 protein, encoded by the *Nfkb1* gene, like the p52 protein (encoded by *Nfkb2*), lacks a transactivation domain. However, it has been proposed to regulate transcription via diverse mechanisms as a subunit of multiple dimeric species. First, the abundant RelA:p50 and c-Rel:p50 heterodimers support transcriptional activation of numerous pro-inflammatory genes, with the RelA and c-Rel subunits providing the transactivation domain[18]. In addition, p50 homodimers have been proposed to serve as transcriptional repressors, due to their absence of a transactivation domain and potential to compete for binding to some NF-κB motifs[19,20]. Interactions between p50 homodimers and histone deacetylases have also been suggested to contribute to transcriptional repression[21]. Finally, p50 has been shown to interact with nuclear inhibitor of NF-κB (IκB) proteins, which serve as co-activators or co-repressors[22–27].

The mouse and human genomes encode eight IκB proteins, three of which - IκBα, IκBβ, and IκBε – play major roles in sequestering NF-κB dimers in the cytoplasm prior to cell stimulation[28]. Two other IκB-like proteins are components of the p105 and p100 precursors to the NF-κB p50 and p52 proteins. Of greatest relevance to this study, the remaining three IκB proteins, IκBζ (*Nfkbiz*), IκBNS (IκBδ, *Nfkbid*), and Bcl3 (*Bcl3*), are found predominantly in the nucleus and each has been shown to interact with the NF-κB p50 protein. IκBζ primarily serves as a transcriptional co-activator, IκBNS is primarily a co-repressor, and Bcl3 has been proposed to contribute to both activation and repression[23,24,29,30]. Notably, other IκBs, including, IκBα and IκBβ, also enter the nucleus and help regulate NF-κB functions.

Although much has been learned, the regulatory logic through which different NF-κB dimeric species and their IκB partners contribute to selective gene regulation remains poorly understood. This knowledge has been difficult to achieve for a variety of reasons, including

considerable redundancy between dimeric species, the similarity of binding motifs for many NF-κB dimers, and the complexity of the mutant phenotypes.

To gain further insight, we performed nascent transcript RNA-seq with bone marrow-derived macrophages (BMDMs) from wild-type (WT), *Nfkb1*$^{-/-}$, *Nfkbiz*$^{-/-}$, *Nfkbid*$^{-/-}$, and *Bcl3*$^{-/-}$ mice. The study of nascent transcripts focused attention on the impact of each mutation on transcript levels separately from the effects on mRNA stability, which would impact mRNA analyses. Despite the diverse mechanisms by which p50 can regulate gene transcription, we observed strong p50-dependent transcription at only a small number of key inflammatory and immunoregulatory genes, including *Il6*, *Il1b*, *Nos2*, *Adamts4*, and *Lcn2*. We also observed an unanticipated degree of overlap between p50-dependent genes and genes exhibiting strong dependence on IκBζ. ChIP-seq analysis revealed extensive co-occupancy of thousands of genomic sites by p50, IκBζ, and other NF-κB family members. Strong enrichment of p50-IκBζ co-dependent genes was found near a small subset of genomic sites where both IκBζ binding and RelA binding by ChIP-seq are dependent on the presence of p50. Motifs representative of p50 homodimer binding were not enriched near p50-IκBζ co-dependent genes or at genomic locations with p50-dependent IκBζ binding. Instead, in vitro interaction studies, as well as ChIP-seq and sequential ChIP-seq experiments, suggest that IκBζ may interact functionally with p50:RelA heterodimers.

Of particular interest, both IκBζ mRNA and the set of p50-IκBζ-co-dependent genes identified by nascent transcript RNA-seq were among the genes that exhibit the greatest differential expression in response to TLR4 and TNFR signaling. This suggests that poor expression of IκBζ in response to TNFR signaling, partly due to the selective instability of IκBζ mRNA[31,32], represents a third major contributor, in addition to IRF3 and kinetic differences[11,33], to the differential TLR4 and TNFR responses. Adding support to this hypothesis, IκBζ overexpression restored the expression of a subset of IκBζ-dependent genes in response to TNFR signaling.

RESULTS

Gene Dependence on p50

To analyze the underlying logic through which distinct members of the NF-κB family help coordinate selective transcriptional responses, we first performed RNA-seq with nascent chromatin-associated transcripts isolated from BMDMs prepared from WT and *Nfkb1*$^{-/-}$ mice stimulated with Lipid A for 0, 0.5, 1, 2, and 6hr. Relatively short time points were examined to increase the probability of capturing the direct gene targets of the NF-κB p50 protein.

In unstimulated BMDMs, the absence of p50 had little effect on nascent transcript levels (Figure 3-1A, left). Although p50 homodimers have been proposed to contribute to transcriptional repression, these results reveal that its loss in unstimulated cells is insufficient to significantly increase the transcription of any genes. Interestingly, at the 0.5hr time point, only 4 genes were strongly (< 33% relative to WT) and significantly (p-adj < 0.01) dependent on p50 (Figure 3-1A, right). *Egr1* stands out among these four genes in its magnitude of dependence. Importantly, *Egr1* is unlikely to be a p50 target. Instead, this strong and unique dependence is likely due to the dependence of *Egr1* transcriptional activation on the mitogen-activated protein kinase (MAPK), ERK, which activates serum response factor (SRF)-associated ternary complex factors (TCFs) that are critical for *Egr1* transcription[34]. ERK activation is dependent on the MAPK kinase kinase, Tpl2, whose activity is strongly deficient in *Nfkb1*$^{-/-}$ macrophages because Tpl2 stability requires association with the p105 precursor protein encoded by *Nfkb1* gene[35] (i.e. the precursor to p50). RNA-seq analyses with lipid A-stimulated *Tpl2*$^{-/-}$ BMDMs[36], and with lipid A-stimulated WT BMDMs in the presence of an ERK inhibitor (Figure S3-1), demonstrated that *Egr1* is by far the most strongly impacted gene in both settings. These results further strengthen the notion that the *Egr1* deficiency in *Nfkb1*$^{-/-}$ at the 30-min time point is due to the absence of both Tpl2 and activated ERK rather than the absence of p50 (note that ERK and the p38 MAPK activate many other TLR4-induced genes in a redundant manner[37]).

At 1, 2, and 6hr post-stimulation, the transcription of 62 genes (WT RPKM > 3; < 33% expression relative to WT; p-value < 0.01) was strongly diminished in the *Nfkb1*$^{-/-}$ BMDMs. This

includes only 23 of the 236 genes that are most strongly induced (> 10-fold) by lipid A at these time points (Figure 3-1B). Despite the small number of p50-dependent genes, this subset includes key inflammatory and immunoregulatory genes, including *Il6*, *Il1b*, *Nos2*, *Adamts4*, and *Lcn2.* Because of the biological importance of these genes, understanding the reason for their non-redundant requirement for p50 is of considerable interest for an understanding of NF-κB selectivity mechanisms.

Initial Analysis of Nuclear IκBs

The NF-κB p50 protein has the potential to regulate transcription via diverse mechanisms, beginning with its assembly into RelA:p50 heterodimers, which is thought to be the most abundant NF-κB dimeric species in many cell types. One possibility is that only a subset of RelA:p50 targets are dependent on p50 because its closely related paralog, p52, can functionally substitute for p50 at most but not all target genes. Another possibility is that p50-dependence may occur at genes activated by p50 homodimers associated with one or more of the nuclear IκB proteins, IκBζ, Bcl3, and IκBNS.

To explore a possible role for nuclear IκBs at the p50-dependent genes, we performed nascent transcript RNA-seq with BMDMs from *Nfkbiz*$^{-/-}$(IκBζ), *Bcl3*$^{-/-}$, or *Nfkbid*$^{-/-}$ (IκBNS) mice. In all cases, WT and mutant cells were compared in parallel after lipid A stimulation for 0, 0.5, 1, 2, or 6hr. Significant differences were not observed with either *Bcl3*$^{-/-}$, or *Nfkbid*$^{-/-}$ BMDMs in this setting (Figure 3-1C). However, in *Nfkbiz*$^{-/-}$ BMDMs, 134 genes exhibited strongly diminished transcript levels (WT RPKM > 3; < 33% relative to WT; p-value < 0.01), with 30, 62, and 104 of these genes diminished at the 1, 2, and 6hr time points, respectively (Figure 3-1C). No genes were diminished at the 0 and 0.5hr time points, consistent with prior knowledge that *Nfkbiz* is only weakly expressed in unstimulated BMDMs and is a potently induced primary response gene.

Interestingly, a large percentage (38%) of the p50-dependent genes were found to exhibit p50/IκBζ-co-dependence, with 24%, 50%, and 46% of p50-dependent genes at the 1, 2, and 6hr time points also exhibiting IκBζ-dependence at the same time point (Figure 3-1D). Furthermore,

a closer analysis revealed that 85% of the genes exhibiting the strongest dependence on p50 at the 1hr time point (< 12% relative to WT; p-value < 0.01) exhibited IκBζ dependence (Figure 3-1E). Thus, a surprisingly large contribution to p50-dependent gene transcription is via a single mechanism: namely, p50's collaboration with IκBζ. This finding provides initial evidence of a dedicated p50/IκBζ regulatory pathway that may contribute to selective pro-inflammatory responses.

Notably, p50/IκBζ co-dependence was observed at most of the key inflammatory and immunoregulatory genes that exhibited p50 dependence, including *Il6*, *Il1b*, *Nos2*, *Adamts4*, and *Lcn2* (Figure 3-1F). Gene ontology analysis suggested a role for co-dependent genes in Th17 cell development and function. The statistically enriched pathways in Gene Ontology analysis for Biological Processes (GO: BP) and in the Kyoto Encyclopedia of Genes and Genomes (KEGG) are immune system processes and Il-17 signaling, respectively (Figure 3-2A). Consistent with this role, media from *Nfkb1*$^{-/-}$ macrophages exhibited a reduced capacity to support the differentiation of naïve CD4+ T-cells toward a Th17-cell fate (Figure 3-2B). However, the IL-6 and IL-1B cytokines that exhibit strong p50/IκBζ co-dependent genes carry out pleiotropic functions, making it difficult to predict how their selective regulation by a p50/IκBζ regulatory pathway might allow them to properly carry out their functions.

Induction Properties of p50/IκBζ Co-Dependent Genes

To better understand the mechanisms regulating p50/ IκBζ co-dependent genes, as well as genes dependent only on p50 or on IκBζ, we investigated their transcriptional induction kinetics and their dependence on the protein synthesis inhibitor, cycloheximide (CHX). Because *Nfkbiz* is poorly expressed in unstimulated cells but is strongly induced in response to lipid A, we initially expected all IκBζ-dependent and p50/IκBζ-co-dependent genes to be CHX-sensitive secondary response genes induced with delayed kinetics.

Contrary to expectations, the 18 p50/IκBζ-co-dependent genes (< 33% relative to WT; p-value < 0.01) exhibited diverse properties (Figure 3-3A). Several genes, including *Il6*, *Nos2*, and

Batf exhibited properties of conventional secondary response genes, with nascent transcript induction first observed at the 1- and 2-hr time points; with the initial induction strongly diminished in the presence of CHX; and with transcripts at all induced time points diminished in both *Nfkb1*$^{-/-}$ and *Nfkbiz*$^{-/-}$ cells. *Il6* is shown as an example of these genes (Figure 3-3B). In contrast, *Il1b* induction begins with a strong primary response component, characterized by CHX resistance and independence of both p50 and IκBζ at the 0.5-hr time point. However, sustained transcription at 1 and 2hr is dependent on both p50 and IκBζ and exhibits CHX sensitivity (presumably due to IκBζ's low expression prior to its transcriptional induction). A third small group of genes (3 of 18 co-dependent genes) is exemplified by *Blnk* (Figure 3-3B), whose transcript levels are highest in unstimulated cells and are repressed following lipid A stimulation. These genes exhibit p50/IκBζ co-dependence 1 and 2hr post-stimulation because p50 and IκBζ appear to enhance the magnitude of transcriptional repression. All three of these general modes of regulation (with additional variations) were also observed with genes that exhibited dependence on only p50 or IκBζ. *Il12b* and *Dusp2* are shown as examples of genes classified as IκBζ-dependent/p50-independent, and *Cxcl16* is shown as an example of genes classified as p50-dependent/IκBζ-independent. It may be noteworthy that *Dusp2* and *Cxcl16* show some evidence of p50/IκBζ co-dependence but were classified as dependent on only one of the two factors because the magnitude of the impact of the second mutant did not reach the thresholds used for this analysis. Thus, the close collaboration between p50 and IκBζ may be understated. Together, these results show that many p50/IκBζ target genes have the expected secondary response characteristics. However, p50 and IκBζ regulation and co-regulation can also help sustain the expression of primary response genes and of genes repressed by lipid A.

Analysis of *Nfkb2*$^{-/-}$ and *Nfkb1*$^{-/-}$*Nfkb2*$^{-/-}$ Macrophage

As noted above, many NF-κB target genes activated by p50-containing dimers may not exhibit p50-dependent transcription in lipid A-stimulated BMDMs due to redundancy with p50's closely related paralog, p52. Redundancy with p52 may also contribute to the existence of a set of IκBζ-

dependent/p50-independent genes. On the other hand, the extensive overlap between p50-dependent genes and IκBζ-dependent genes shows that p52 cannot compensate for the loss of p50 in the genes defined as p50/IκBζ co-dependent.

To gain insight into the contribution of p52, we first performed RNA-seq (nascent and mRNA) with *Nfkb2*$^{-/-}$ BMDMs. The results revealed a small number of p52-dependent genes. However, none of these genes exhibited IκBζ dependence (data not shown). This result demonstrates the absence of genes that are uniquely p52/IκBζ co-dependent.

To examine possible redundancy between p50 and p52, we created a J2 retrovirus-immortalized macrophage line from *Nfkb1*$^{-/-}$ mice and then used CRISPR-Cas9 editing to disrupt the *Nfkb2* gene in this line. RNA-seq analysis (mRNA) comparing the *Nfkb1*$^{-/-}$*Nfkb2*$^{-/-}$ line and the parental *Nfkb1*$^{-/-}$ line to a WT J2-transformed macrophage line revealed that the combined p50 and p52 deficiency yielded reduced transcript levels at a much larger number of inducible genes than observed in the absence of p50 alone. This confirms the existence of considerable redundancy between the two paralogs. However, 71% of strongly induced genes remained induced, either because they can be activated by NF-κB dimers that do not include either p50 or p52 or because their induction is entirely NF-κB independent (data not shown).

Focusing on IκBζ-dependent genes, we found that only 2 additional IκBζ-dependent/p50-independent genes exhibited reduced transcript levels in the *Nfkb1*$^{-/-}$*Nfkb2*$^{-/-}$ line. This result demonstrates that IκBζ-dependent/p50-independent genes are not compensated for by the presence of *Nfkb2* but rather arise from another unknown mechanism.

Initial p50, IκBζ, and RelA Chip-Seq Analyses

To elucidate the mechanism underlying the regulation of p50/IκBζ co-dependent genes, we performed ChIP-seq with p50, RelA, and IκBζ antibodies in BMDMs stimulated with lipid A for 0, 0.5, 1, and 2hr. We observed 2,311, 6,123, and 3,310 reproducible peaks (peak score >19 and RPKM >3 in 2/2 biological replicates at one or more time points) with the p50, RelA, and IκBζ antibodies, respectively. We are unable to determine whether the different peak numbers are due

to different antibody qualities or different numbers of genomic interactions. Despite the different peak numbers, strong correlations were observed for pair-wise comparisons of p50, RelA, and IκBζ ChIP-seq data sets (Figure 3-4B). These results demonstrate that p50- and IκBζ-dependent transcription is not simply due to highly selective genomic interactions of these proteins, as the number of genomic interactions dramatically exceeds the number of functionally dependent genes.

To better understand the relationships between the three proteins, we compared peak numbers at each time point (Figure 3-4A). This reveals a large number of p50 and RelA peaks at the 1- and 2-hr time points. IκBζ binding is almost exclusively found at the 2-hr time point, which aligns with the kinetic profile of *Nfkbiz* induction. Notably, p50 peaks were also prevalent in unstimulated cells, consistent with prior knowledge of p50 homodimer localization to the nucleus prior to cell stimulation.

Next, k-means cluster analysis with all p50, RelA, and IκBζ binding sites throughout the time-course allowed the peaks observed for each protein to be classified into six distinct kinetic clusters (Figure 3-4C). p50 peaks were prevalent in clusters 1-5 and RelA peaks were prevalent in clusters 2-5, but both p50 and RelA peaks were rarely assigned to cluster 6. In contrast, 99.9% of the IκBζ peaks were assigned to cluster 6, which is characterized by maximum binding at the 2-hr time point and much smaller peak scores at 0.5 and 1hr. Because a high percentage of IκBζ peaks coincide with p50 and RelA peaks, these results suggest that IκBζ associates with genomic locations at which an NF-κB dimer is already bound. Although the possibility of NF-κB dimer exchange coinciding with the late IκBζ binding cannot be excluded.

Finally, we used ATAC-seq to examine the chromatin state of genomic sites bound by p50, RelA, and IκBζ. In kinetic clusters in which p50 binding first occurred at early time points, the genomic locations usually exhibited open chromatin as demonstrated by ATAC-seq signals in unstimulated cells (Figure 3-4D). In contrast, in cluster 6, representative of locations at which p50 binding was delayed, ATAC-seq signals were not observed in unstimulated cells, suggesting the need for inducible nucleosome remodeling prior to remodeling (Figure 3-4D). Although IκBζ binding typically occurred late, the genomic locations of its binding were typically assembled in

open chromatin prior to its binding, consistent with the evidence that IκBζ- joins a pre-bound NF-κB dimer at these locations (Figure 3-4D).

Relationship between IκBζ Binding and p50 versus RelA Preferential Binding

The above results show that p50, RelA, and IκBζ can interact with thousands of genomic locations, with extensive overlap. The results further suggest that IκBζ typically binds NF-κB dimers that are already bound to DNA, with the genomic sites usually assembled into open chromatin before IκBζ binding. However, the above results do not reveal distinguishing characteristics of the relatively small number of genomic regions at which IκBζ binding appears to be functionally important, and the results do not reveal why these genomic sites and their corresponding target genes require IκBζ for activity.

To gain insight into these questions, we asked whether IκBζ binding exhibits a preference for genomic sites preferentially bound by p50 in comparison to RelA. For this analysis, we calculated the RelA and p50 RPKM ratio at each genomic site containing a called peak for either protein (Figure 3-5A). Notably, only about 2% of peaks exhibit strong p50 or RelA preference. We then separated the peaks into bins based on this ratio and examined the prevalence of IκBζ binding within each bin. This analysis reveals that IκBζ is greatly enriched at sites co-bound by p50 and RelA (p50/RelA ratio of 0.3 – 1.6), but quickly declines in bins with strong p50 or RelA preference (Figure 3-5B). This indicates that IκBζ does not have a preference for binding with p50 homodimers, which would have elevated p50/RelA ratios.

Notably, motif enrichment analysis across the spectrum of p50-RelA binding ratios revealed that bins containing peaks with preferential p50 binding were enriched for NF-κB motifs representative of p50 homodimer binding, with three G:C-bps in each half-site. Bins containing peaks with preferential RelA binding were enriched for motifs representative of RelA homodimer binding, with two G:C-bps in each half-site (Figure 3-5C). A comparison between p50 binding and binding of the c-Rel member of the NF-κB family (performed with c-Rel ChIP-seq datasets) revealed similar results to the p50-RelA comparison (Figure 3-5D). An example of a p50-

preferential binding site is found at the *H2-Aa* promoter (Figure 3-5F). At the majority of p50-preferential binding sites, p50 is strongly bound prior to stimulation (data not shown). Although we did not see any genes with strong enhancement in the absence of p50 (Figure 3-1A), we can observe a depletion of strongly induced (FC > 5) genes near p50-preferential sites (Figure 3-5E). This result suggests that sites preferential bound by p50 (possibly p50 homodimer binding sites) may rarely contribute to lipid A-induced transcription.

Relationship between p50/IκBζ-Co-Dependent Genes, IκBζ- and p50-Dependent Genes, and p50, RelA, and IκBζ Binding

We next asked whether p50/IκBζ-co-dependent, IκBζ-dependent genes, or p50-dependent genes correlate either with genomic regions that exhibit preferential p50 binding or with NF-κB motifs known to preferentially bind p50 homodimers. For this analysis, each peak was annotated to its closest gene within 5 kb. The prevalence of peaks annotating to p50/IκBζ-, IκBζ-, and p50-dependent genes was analyzed across the spectrum of p50-RelA RPKM ratios. Peaks annotating to p50-IκBζ-co-dependent genes, p50-dependent genes (Figure 3-5E), or IκBζ-dependent genes (data not shown) showed no significant bias toward either p50-preferential peaks or RelA preferential peaks. In addition, motif enrichment analysis showed that peaks linked to p50/IκBζ-co-dependent and p50-dependent genes showed no preference toward motifs representative of p50 homodimer binding (data not shown). These results further suggest that p50 homodimers may not play a role in the modulation of p50/IκBζ co-dependent or p50-dependent genes.

Because the above approaches failed to uncover a property that might distinguish functional IκBζ interactions and failed to show a relationship between IκBζ binding and p50-preferential binding, we extended our analysis by identifying IκBζ binding events that are dependent on the presence of p50. For this analysis, IκBζ ChIP-seq was performed with WT and *Nfkb1*$^{-/-}$ BMDMs. Surprisingly, only about 4% of IκBζ genomic interactions exhibited a strong dependence (< 33% binding of IκBζ in *Nfkb1*$^{-/-}$ relative to WT) on p50, with the remaining IκBζ peaks exhibiting either weak or no dependence on p50 (Figure 3-6A).

To evaluate the possible significance of this small set of p50-dependent IκBζ binding events, we examined their locations in relation to p50/IκBζ-, IκBζ-, and p50-dependent genes. Interestingly, this analysis revealed a high frequency IκBζ ChIP-seq peaks that annotate to p50/IκBζ co-dependent or p50-dependent genes among the small fraction of IκBζ ChIP-seq peaks that require p50 (Figure 3-6A). Peaks annotating to IκBζ-dependent genes did not show this same relationship (data not shown). To quantify this finding, all IκBζ ChIP-seq peaks (n=3,733) were separated into 15 equal bins on the basis of their quantitative dependence on p50. The percentage of peaks within each bin that annotates to p50/IκBζ co-dependent or p50-dependent genes was then calculated. In the bin containing IκBζ peaks that exhibit the strongest requirement for p50, 15% annotate to p50/IκBζ co-dependent genes, with another 17% of these peaks annotating to p50-dependent genes (Figure 3-6B). Furthermore, of the 25 p50/IκBζ co-dependent genes (Figure 3-1F), 17 (68%) have a p50-dependent IκBζ binding event within 30 kb of the TSS (Figure 3-6C). These results demonstrate that, although IκBζ genomic interactions can be observed by ChIP-seq at thousands of genomic locations, we can greatly enrich for those with putative functional relevance by examining the dependence of IκBζ binding on p50.

The above results suggest that p50-dependence of IκBζ binding may be predictive of a more specific interaction or a relatively rare binding conformation that supports IκBζ function (see Discussion). To explore these possibilities, we examined binding strength, binding preferences, H3K27ac marks, and binding motif differences between p50-dependent and p50-independent genomic interaction sites. First, IκBζ, RelA, and p50 peak scores were comparable at p50-dependent and p50-independent IκBζ binding sites. In addition, no difference in p50 or RelA preferential binding was observed at p50-dependent versus p50-independent IκBζ binding sites (data not shown). To compare motif enrichment at p50-dependent versus p50-independent IκBζ binding sites, we binned all IκBζ binding sites based on their magnitude of p50 dependence. We then examined the enrichment of known NF-κB motifs in each bin. We find the strongest enrichment of NF-κB motifs at p50-dependent IκBζ binding sites (Figure 3-6D). This further strengthens the idea that IκBζ interacts with NF-κB dimers at these locations but does not

distinguish these binding sites from the thousands of other NF-κB bound sites in the genome. The relatively low prevalence of NF-κB motifs at p50-independent IκBζ peaks raises the possibility that IκBζ binding to these sites often does not require an interaction with an NF-κB dimer. However, NF-κB co-occupies a high percentage of these sites with IκBζ (data not shown), suggesting instead that both IκBζ and an NF-κB dimer occupy these sites, often in the absence of an NF-κB consensus motif.

Finally, we explored the relationship between IκBζ binding dependence on p50 and H3K37ac, a histone modification frequently used as a marker of a transcriptionally active state. At a genome-wide scale, the kinetics of IκBζ binding does not correlate with the kinetics of the histone H3K27ac signal. However, at IκBζ peaks that annotate to p50/IκBζ co-dependent genes, the kinetics of IκBζ correlates closely with the kinetics of the H3K27ac modification, even though the sites usually exhibit open chromatin prior to stimulation and are bound by p50 and RelA at earlier time points. An Il6 enhancer provides one example of these kinetic relationships (Figure 3-6E). This finding suggests that, although an NF-κB dimer can bind earlier, this binding is insufficient to activate the control region. Thus, a kinetic relationship between IκBζ binding and the H3K27ac modification may be another feature of functionally important IκBζ binding events, along with the p50-dependence of IκBζ binding.

In Vivo and In Vitro Analysis of IκBζ Interactions with p50 and RelA

The above results demonstrate that IκBζ co-occupies thousands of genomic sites with the NF-κB p50 and RelA subunits, with no bias toward sites exhibiting p50 preferential binding. Most prior studies of IκBζ have proposed that it functions primarily by contributing a transactivation domain to p50 homodimers. However, a few studies have suggested that it may also interact with RelA:p50 heterodimers[38].

An examination of our ChIP-seq data added further uncertainty to the question of which NF-κB dimer carries out functional interactions with IκBζ. Specifically, genomic sites exhibiting p50-dependent IκBζ binding in close proximity to p50/IκBζ-co-dependent genes showed ChIP-

seq peaks for RelA in addition to peaks for p50 and IκBζ. The underlying NF-κB motifs at these locations typically resembled a RelA:p50 heterodimer motif, with three G:C bps in one half-site and two G:C bps in the other.

To further explore this issue, we performed RelA ChIP-seq in *Nfkb1*$^{-/-}$ BMDMs. The results revealed that at 76% of p50-dependent IκBζ binding sites, RelA binding was also p50-dependent (Figure 3-7A; see also Figure 3-7C). This finding supports the possibility that IκBζ binds a RelA:p50 heterodimer, consistent with the presence of ChIP-seq peaks for all three proteins. Alternatively, exchange between RelA:p50 heterodimers and p50 homodimers may occur at these locations, with IκBζ associated only with alleles bound by p50 homodimers.

To begin to distinguish between these possibilities, we first performed co-immunoprecipitation experiments to examine IκBζ's association with RelA in nuclear extracts from a J2-transformed BMDM cell line. Following immunoprecipitation with RelA antibodies, immunoblots were performed to monitor the co-immunoprecipitation of the p50, RelA, and IκBζ proteins. Importantly, titrations of the RelA, p50, and IκBζ input proteins were included to allow us to calculate the approximate percentage of the proteins co-immunoprecipitated. The results revealed that IκBζ can co-immunoprecipitate with RelA antibodies. With 19%, 9%, and 4% of total RelA, p50, and IκBζ, respectively, co-immunoprecipitating with RelA (Figure 3-7D). The finding that, in co-immunoprecipitation experiments, only twice as many p50 molecules are associated with RelA than IκBζ molecules adds support to the possibility that IκBζ can associate with RelA:p50 heterodimers in vivo. Parallel experiments with p50 antibodies resulting in the co-immunoprecipitation of approximately 3% of the IκBζ molecules in the sample (data not shown), confirming the expected p50- IκBζ interaction.

We also performed electrophoretic mobility shift assays (EMSA) with extracts containing overexpressed RelA, p50, and IκBζ, either expressed individually and combined or co-expressed to determine whether complexes containing IκBζ associated with DNA-bound p50 homodimers or RelA:p50 heterodimers can be detected. However, as observed previously[39], IκBζ strongly inhibited DNA binding by both p50 homodimer and RelA:p50 heterodimers (data not shown). The

inhibition of DNA-binding provides further evidence that IκBζ can bind both RelA:p50 heterodimers and p50 homodimers in vitro. However, in vivo, IκBζ co-occupies DNA with NF-κB dimers. One possible reason for the in vitro inhibition is that IκBζ requires processing or a post-translational modification for co-binding DNA with NF-κB dimers, with the processing event or modification absent in extracts containing overexpressed protein.

Finally, to further examine which NF-κB dimers can support IκBζ's association with DNA in vivo, we employed Sequential-ChIP-seq experiments. For these experiments, we first performed ChIP with RelA antibodies. We then used 10mM DTT to dissociate the immunoprecipitated chromatin from the antibody-bead complex, followed by a second immunoprecipitation with antibodies directed against either p50, IκBζ, or p53 (a negative control). The results revealed that both p50 and IκBζ are often present in chromatin fragments associated with RelA, supporting a model in which IκBζ is associated with NF-κB dimers containing RelA, most likely RelA:p50 heterodimers (Figures 3-7B, C). Additional studies will be needed to understand the structural basis of IκBζ's interactions with NF-κB dimers and why these interactions are required for the activity of a select subset of transcriptional regulatory regions.

Role of IκBζ in the Differential Responses to TLR4 and TNFR Signaling

As described above, a small but diverse set of key inflammatory and immunoregulatory genes exhibit strong p50/IκBζ co-dependence. Other key genes, including *Il12b*, exhibit strong dependence on one or both factors in lipid A-stimulated macrophages. Gene ontology analysis suggested a role for the subset of co-dependent genes in Th17 differentiation (Figure 3-2A). However, the co-dependent genes have diverse functions, potentially allowing the p50/IκBζ pathway to impact multiple biological processes.

To expand our knowledge of the potential biological significance of the p50/IκBζ regulatory pathway, we searched for settings in which IκBζ might contribute to differential gene transcription. This effort led to a focus on BMDMs stimulated with lipid A versus TNF. Although *Nfkbiz* is a

primary response gene that is potently activated by lipid A, previous studies have shown that the *Nfkbiz* gene is poorly expressed in TNF-stimulated macrophages[30,32].

To extend the characterization of *Nfkbiz* differential expression, we performed nascent transcript RNA-seq and mRNA-seq with TNF-stimulated BMDMs. A comparative analysis revealed a surprisingly large magnitude of *Nfkbiz* differential expression between TNF and lipid A stimulation. *Nfkbiz* nascent transcripts were found to be 18-fold more abundant after lipid A stimulation than after TNF stimulation at the 1.0h time point. In addition, *Nfkbiz* mRNA was 62-fold more abundant following lipid A stimulation than TNF stimulation (Figure 3-8A). The greater magnitude of the mRNA difference is likely due to a previously described[31,32] instability of the *Nfkbiz* mRNA following TNF stimulation, suggesting the existence of a mechanism to ensure that IκBζ levels remain extremely low following TNF stimulation.

The large magnitude of *Nfkbiz* differential expression was highly unusual, as revealed by a comparison of *Nfkbiz* to the 10 other genes that comprise the NF-κB family, and the IκB family, all of which are induced by inflammatory stimuli. This comparison revealed that *Nfkbiz* mRNA transcripts exhibit a differential expression magnitude that is much greater than that of all other NF-κB and IκB members (Figure 3-8B). Furthermore, among the 18 transcriptional regulators activated during the primary response to lipid A or TNF, *Nfkbiz* exhibits the largest differential expression at the 1h time point (data not shown). Finally, among all strongly induced (FC > 10) primary response genes activated by lipid A or TNF, *Nfkbiz* ranks 12th among 132 primary response genes (Figure 3-8C). These results suggest that both transcriptional and post-transcriptional mechanisms have emerged to ensure that the abundance of IκBζ remains extremely low following TNF stimulation.

Consistent with the strongly differential expression of *Nfkbiz*, all (25 or 25) of the p50/IκBζ co-dependent and 81% of the IκBζ-dependent genes exhibited strong differential nascent transcript levels (< 33% TNF-stimulated RPKM relative to lipid A-stimulated RPKM) in a comparison of lipid A and TNF-stimulated BMDMs (Figure 3-8D). In fact, among the genes showing the strongest differential expression (TNF/ lipid A percent expression < 2%), 41% are

among the small group of p50/ IκBζ or IκBζ targets. Moreover, these genes are not dependent on IRF3 (from an analysis of lipid A-stimulated *Irf3*$^{-/-}$ BMDMs), demonstrating that their low expression is not due to the absence of IRF3 activation (Figure 3-8D).

To estimate the relative contributions of IRF3 and IκBζ to the differential responses of BMDMs to lipid A versus TNF, we used nascent transcript RNA-seq data sets from *Ifnar*$^{-/-}$ and *Irf3*$^{-/-}$ BMDMs to determine the number of genes that exhibit strong (< 33% relative to WT) dependence on IRF3 or its target, *Ifnb1*. This analysis revealed that 10.9% of lipid A versus TNF differentially expressed genes exhibit IRF3/IFNAR-dependence. In comparison, 11.7% of lipid A versus TNF differentially expressed genes exhibit IκBζ-dependence. Focusing more closely on the 27 genes that exhibit the strongest lipid A/TNF differential expression (nascent transcript level in response to TNF < 2% of the level in response to lipid A), we found that IRF3/IFNAR-dependence can account for the differential expression of 8 (29.6%) of these genes. In contrast, 11 (40.7%) of these genes exhibit IκBζ-dependence (Figure 3-7D).

To determine whether an increased abundance of IκBζ is sufficient to elevate transcript levels for IκBζ-dependent genes in response to TNF stimulation, we used retroviral transduction to overexpress IκBζ in BMDMs, followed by stimulation with lipid A or TNF. In the presence of overexpressed IκBζ, the induction magnitude of two IκBζ-dependent genes, *Lcn2* and *Tapap*, was comparable after TNF stimulation to the induction magnitude observed after lipid A stimulation, showing efficient rescue by overexpressed IκBζ (Figure 3-8E). In addition, the magnitude of the lipid A/TNF differential induction of *Il1b* was reduced in the presence of overexpressed IκBζ (Figure 3-8E). However, the lipid A/TNF differential induction of other p50/IκBζ co-dependent and IκBζ-dependent genes was not impacted by IκBζ overexpression (Figure 3-8E). One possible reason for the limited impact of IκBζ overexpression is that IκBζ activity may require complex regulatory mechanisms that cannot be achieved with the overexpressed protein. The existence of such mechanisms is supported by the inability of recombinant IκBζ to co-occupy DNA with NF-κB dimers in vitro (see above). Nevertheless, we

cannot exclude the alternative possibility that many IκBζ targets are poorly expressed in response to TNF for multiple reasons.

DISCUSSION

In this study, we performed a genomics-centric analysis of selective transcriptional regulation in stimulated BMDMs by the NF-κB p50 protein and the nuclear IκB protein, IκBζ. The study began with the identification of a small group of key inflammatory and immunoregulatory genes that exhibited a strong dependence on p50 in nascent transcript RNA-seq experiments performed with lipid A-stimulated BMDMs. Strong overlap between p50-dependent and IκBζ-dependent genes revealed that a defined p50/IκBζ pathway makes a major contribution to the regulation of this key set of genes. Although p50, RelA, and IκBζ occupy thousands of genomic sites, a potentially defining characteristic of functionally meaningful binding events was the p50-dependence of IκBζ binding at a small subset of sites. Similar temporal kinetics of IκBζ binding and H3K27ac deposition also appeared to distinguish functional from non-functional IκBζ interactions. Sequential ChIP-seq and biochemical results provided strong support for the possibility that IκBζ functionally interacts with RelA:p50 heterodimers rather than p50 homodimers. Biologically, the p50/IκBζ pathway contributes to the selective regulation of key inflammatory and immunoregulatory genes, some with pleiotropic functions. Moreover, this pathway appears to make a major contribution to the differential responses of macrophages to the microbial product, lipid A, and the inflammatory cytokine, TNF.

The small number of p50-dependent genes and the strong overlap with IκBζ-dependent genes was unexpected, given that p50 also contributes to stimulus-induced transcription as a component of the abundant RelA:p50 and c-Rel:p50 heterodimers. Redundancy between p50 and its closely related paralog, p52, provides a partial explanation for the limited number of p50-dependent genes, as a larger number of lipid A-induced genes exhibited reduced transcription in an *Nfkb1*$^{-/-}$*Nfkb2*$^{-/-}$ macrophage line. However, lipid A induction of a large number of genes was retained in these cells, suggesting that either a large number of lipid A-induced genes are NF-κB

independent or that other NF-κB dimers, such as c-Rel homodimers, RelA homodimers, and c-Rel:RelA heterodimers (as well as RelB-containing dimers) are sufficient for the induction of many genes. Importantly, p52 and other NF-κB family members are unable to compensate for the loss of p50 at a high percentage of IκBζ-dependent genes, showing an unusually close relationship between p50 and IκBζ.

Prior studies have shown that IκBζ binds p50, but the fact that IκBζ dependence characterizes a high percentage of genes exhibiting a non-redundant dependence on p50 was not previously known. Most prior studies of IκBζ have suggested that its function involves association with p50 homodimers, with IκBζ providing an activation domain to a homodimer that otherwise would lack such a domain. However, our sequential ChIP-seq and biochemical results support a smaller number of prior studies that revealed in vitro interactions between IκBζ and RelA:p50 heterodimers. If RelA:p50 heterodimers carry out functional interactions with IκBζ, a mechanistic understanding of IκBζ's role would require further exploration because RelA:p50 heterodimers possess RelA's transactivation domain. According to combinatorial principles of gene regulation, the transcriptional activation of most if not all genes is thought to require multiple transcription factors, many with transactivation domains, and little is currently known about the mechanisms underlying the requirement for multiple transactivation domains for transcriptional activation. Thus, mechanistically, it is not unreasonable to envision a critical requirement for IκBζ for transcriptional activation of a subset of genes when associated with a RelA:p50 heterodimer.

In addition to the need for further studies of IκBζ's contribution to the activation of its defined target genes, studies are needed to understand why IκBζ's interactions at the vast majority of its genomic binding sites have no apparent functional consequences. One finding that may be relevant to this issue is that most IκBζ genomic interactions are not p50-dependent. This finding suggests that IκBζ and its associated NF-κB dimer may adopt a specific conformation at a limited subset of sites that is critical for its transcriptional activation function.

Another unanswered question is why recombinant IκBζ appears to be incapable of binding DNA-associated p50 homodimers or RelA:p50 heterodimers in EMSA experiments, despite

compelling evidence that it can associate with DNA-bound NF-κB dimers in vivo. This finding, observed by us and others[40], suggests that IκBζ may require a post-translational modification, a processing event, or the presence of another protein to allow association with DNA-bound NF-κB, raising the possibility of another layer of regulation of the p50/IκBζ pathway.

Although our data revealed a strong overlap between p50-dependent and IκBζ-dependent genes, a substantial number of genes exhibited a strong dependence on only one of these two proteins. A subset of these genes exhibits substantial dependence on both proteins but was not classified as co-dependent because they did not meet the stringent thresholds used in our study to define dependence. Nevertheless, some genes clearly exhibit strong dependence on one protein, with no impact of the other factor in loss-of-function experiments. A subset of IκBζ-dependent/p50-independent genes can be attributed to redundancy between p50 and p52, but it is not known whether other IκBζ-dependent/p50-independent genes collaborate with other NF-κB dimers or perhaps contribute to gene regulation in concert with a different transcription factor family. p50-dependent/IκBζ-independent transcription is likely to reflect a requirement at a small set of genes for a p50 dimeric species that does not require IκBζ for its function.

The key immunoregulatory genes that exhibit p50/IκBζ co-dependence, including *Il6* and *Il1b*, are potently activated in a large number of cell types in response to diverse stimuli. We cannot exclude the possibility that p50 and IκBζ are universally required for the activation of these genes, but many genes are regulated by different factors in different biological settings. Although a broad range of abnormalities has been reported in *Nfkb1*$^{-/-}$ and *Nfkb2*$^{-/-}$ mice, even more abnormalities would likely be observed if p50 and IκBζ were universally required for the transcription of the target genes identified in this study.

Finally, the finding that the *Nfkbiz* gene and most IκBζ target genes stand out due to their very large magnitude of differential expression between lipid A- and TNF-stimulated cells highlights the potential biological importance of the differential expression of these genes. The unusually low expression of *Nfkbiz* in TNF-stimulated macrophages appears to be due to the combined influence of a transcriptional mechanism that is unique among primary response genes

and an mRNA stability mechanism that also appears unusual or unique. The well-documented importance of many IκBζ target genes for anti-microbial responses provides justification for their potent activation by lipid A, but the reason transcriptional and post-transcriptional mechanisms evolved to ensure that these genes remain largely silent in response to TNF signaling is less obvious. Nevertheless, the large percentage of IκBζ-dependent genes that exhibit differential expression in response to lipid A versus TNF, and the unusually large magnitudes of their differential expression in comparison to the magnitudes observed with IRF3-dependent genes, suggests that IκBζ and the p50/IκBζ pathway are major contributors to differential lipid A/TNF gene induction in macrophages.

EXPERIMENTAL PROCEDURES

Mice

The *Nfkb1*$^{-/-}$ and *Nfkb2*$^{-/-}$ mice were a gift from Alexander Hoffmann at UCLA. The *Nfkbiz*$^{-/-}$, *Bcl3*$^{-/-}$, and *Nfkbid*$^{-/-}$ mice were kindly provided by Giorgio Trinchieri, M.D. (Laboratory of Experimental Immunology, NIH), Ulrich Siebenlist (Laboratory of Molecular Immunology, NIH) and Ingo Schmitz (Center for Infection Research, Helmholtz), respectively.

BMDM Isolation, Differentiation, and Stimulation

Bone marrow extraction was performed from male mice aged 8-12 weeks. Differentiation into BMDMs was performed as previously described[41] (Purbey et al., Cell 2017). Macrophages were treated on Day 6 with 100ng/mL Lipid A (Sigma) or 10ng/mL TNF (Bio-Tech). For MAPK inhibitor-treated samples, macrophages were pretreated with 1uM of ERK-inhibitor (PD 0325901) for 1.0h prior to stimulation.

RNA-seq

Approximately 5 million BMDMs per sample were stimulated and harvested on day 6 of differentiation. Chromatin-associated RNA and mRNA samples were prepared as previously described (Tong et al., Cell 2016). For chromatin-associated RNA-seq, 60 ng (mRNA-seq 500ng) of RNA was used to generate strand-specific libraries with the TruSeq stranded RNA Kit (Illumia). The library preparation protocol was adapted for chromatin-associated RNA-seq to capture all RNA from the chromatin sample. Libraries were sequenced using single-end (50bps) Illumina Hi-Seq 3000. Reads were aligned to the NCBI37/mm9 genome using Hisat2. RPKMs were calculated as previously described (Tong et al., 2016).

ChIP-Seq and Sequential ChIP-seq

For ChIP-seq experiments, approximately 15 million (20 million for sequential ChIP-seq) macrophages were harvested on day 6 of differentiation. RelA (Cell Signaling, 8242S), p50 (Cell Signaling, 13586), IκBζ (4301779, Sigma), cRel (67489, Cell Signaling) and H3K27ac (Active Motif 39133) antibodies were used. For sequential ChIP-seq, after the first immuno-precipitation, the protein-DNA complex was eluted from the antibody-bead complex with 10mM DTT at 37° C for 30 mins. The eluent was then diluted and the second antibody was added. Then the ChIP protocol proceeds as previously described (Barish et al., 2010). Reads were aligned to the NCBI37/mm9 genome using Hista2. Peak scores and RPKM were calculated as previously described (Tong et al., 2016). Only reproducible peaks with described criteria were maintained for downstream analysis.

ATAC-seq

ATAC-seq libraries were prepared as previously described (Tong et al., 2016; Buenrostro et al., 2015). Reads were mapped to the NCBI37/mm9 genome using Hisat2. Reads were processed and normalized as previously described (Tong et al., 2016).

CRISPR/Cas9 Mutagenesis

Synthetic guide RNAs (gRNAs) were designed using CRISPOR (*http://crispor.tefor.net/*) and Massachusetts Institute of Technology CRISPR Designer (http://crispr.mit.edu). Recombinant Cas9 (Synthego) in combination with gRNAs were electroporated into J2-transformed *Nfkb1*$^{-/-}$ macrophages to delete exon 2 of *Nfkb2.* Single-cell clones were screened by genotyping and confirmed with the sequencing of the *Nfkb1* exon 2.

Motif Analysis

Position weight matrices for NF-κB from the Jaspar motif database were used in Homer motif software (*http://homer.ucsd.edu/homer/*). The 400bps surrounding the center of p50, RelA, or IκBζ ChIP-seq peaks were used in motif analyses.

Gene Ontology

The program gProfiler (https://biit.cs.ut.ee/gprofiler/) was used to search for the enrichment of genes in known biological pathways including gene ontology molecular function, cellular components, biological processes, and KEGG databases.

IκBζ Retroviral Overexpression

The IκBζ expression construct was prepared in a pMSCV vector by VectorBuilder. The construct was verified with DNA sequencing. Viral production was carried out by VectorBuilder and prepared to > 10^7 TU/mL. Retroviral transductions were done as previously described (Sanjabi et al., 2005).

FIGURE LEGENDS:

Figure 3-1. Nascent RNA-seq Analysis of *Nfkb1*$^{-/-}$, *Nfkbiz*$^{-/-}$, *Blc3*$^{-/-}$, and *Nfkbid*$^{-/-}$ BMDMs

(A) Volcano plots are shown of *Nfkb1*$^{-/-}$ versus WT macrophages, either unstimulated (left) and 0.5h lipid A stimulated (right). Light gray dashed lines are shown at 33% or 333% expression

relative to WT (vertical lines) and at a p-value of 0.01 (horizontal line). *Nfkb1*-dependent genes (< 33% relative to WT; p-value < 0.01) are highlighted in yellow. All genes with an RPKM greater than 3 at any time point are included in the analysis. Select genes are labeled and their WT fold-induction is shown at that time point.

(B) A composite volcano plot is shown of all genes with dependence (< 33% relative to WT; p-value < 0.01) on *Nfkb1* from 1hr – 6hr stimulation with lipid A. The minimum percent expression in *Nfkb1*$^{-/-}$ macrophages relative to WT expression from 1hr – 6hr (x-axis) and the corresponding p-value (y-axis) are plotted. Genes in this analysis are expressed with a WT RPKM greater than 3 and an induction greater than 10, at any time from 1hr – 6hr.

(C) The total number of *Nfkb1*$^{-/-}$, *Nfkbiz*$^{-/-}$, *Blc3*$^{-/-}$, or *Nfkbid*$^{-/-}$ dependent genes (< 33% relative to WT; p-value < 0.01) at each timepoint is shown. Genes must have an expression at any time point greater than 3 RPKM.

(D) Venn diagrams reveal the overlap between *Nfkb1*-dependent genes (< 33% relative to WT; p-value < 0.01) and *Nfkbiz*-dependent genes (< 33% relative to WT; p-value < 0.01) at each time point of lipid A stimulation.

(E) A volcano plot of WT versus *Nfkb1*$^{-/-}$ macrophages stimulated with lipid A for 1.0h. Co-dependent genes (< 33% relative to WT; p-value < 0.01 for both *Nfkb1*$^{-/-}$ and *Nfkbiz*$^{-/-}$) are shown in red. The darker dashed line indicates where the percent expression relative to WT is 10% in *Nfkb1*$^{-/-}$.

(F) All *Nfkb1*-dependent genes (< 33% relative to WT; padj < 0.01) or *Nfkbiz*-dependent genes (< 33% relative to WT; padj < 0.01) at any timepoint (0 – 6hr) are graphed. *Nfkb1*-dependent, *Nfkbiz*-dependent, or co-dependent genes are highlighted in orange, blue, or red, respectively.

Figure 3-2. Gene Ontology and Biological Significance of *Nfkb1*-Dependent Genes

(A) Gene ontology analysis of the 62 *Nfkb1*-dependent genes from 0hr – 6hr. The top pathways that emerge from Gene Ontology of Biological Processes (GO: BP) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) are involved in immune system responses and Il-17 signaling,

respectively. The *Nfkb1*-co-dependent genes that classify in these pathways are shown (right). All genes in the Il17-signaling pathway are co-dependent on *Nfkbiz*.

(B) An Il17a/f ELISA from T-cells stimulated with BMDM supernatant, anti-CD3 antibodies, and anti-CD28 antibodies for 24 hrs was performed. The added BMDM supernatant is either from WT or *Nfkb1*$^{-/-}$ BMDMs stimulated with LPS for 24 hours.

Figure 3-3. Kinetic Analysis of Nascent RNA-seq from WT, *Nfkb1*$^{-/-}$, or *Nfkbiz*$^{-/-}$ BMDMs

(A) Kinetics of activation are shown (% max RPKM) for all 104 genes with dependence on *Nfkb1* or *Nfkbiz* (< 33% relative to WT; p-value < 0.01) during the lipid A stimulation time course (0hr – 2hr). The genes are classified based on their expression in CHX-treated samples (red). The order indicates *Nfkb1*/*Nfkbiz* co-dependent secondary response genes (cluster 1; < 33% in both *Nfkb1*$^{-/-}$ and *Nfkbiz*$^{-/-}$; < 33% in CHX treated), *Nfkb1*/*Nfkbiz* co-dependent primary response genes (cluster 2; < 33% in both *Nfkb1*$^{-/-}$and *Nfkbiz*$^{-/-}$; > 33% in CHX treated); *Nfkb1*-dependent secondary response genes (cluster 3; < 33% in *Nfkb1*$^{-/-}$; < 33% in CHX treated), *Nfkb1*-dependent primary response genes (cluster 4; < 33% in *Nfkb1*$^{-/-}$; > 33% in CHX treated), *Nfkbiz*-dependent secondary response genes (cluster 5; < 33% in *Nfkbiz*$^{-/-}$; < 33% in CHX treated), *Nfkbiz*-dependent primary response genes (cluster 6; < 33% in *Nfkbiz*$^{-/-}$; > 33% in CHX treated). Properties of each gene are shown, including the kinetics of dependence on *Nfkb1* or *Nfkbiz* (gray); the minimum percent expression in *Nfkb1*$^{-/-}$ (green), *Nfkbiz*$^{-/-}$ (blue), or CHX-treated (red); and the fold induction (purple).

(B) Kinetic plots of select representative genes from WT, *Nfkb1*$^{-/-}$, or *Nfkbiz*$^{-/-}$ macrophages stimulated with lipid A for 0, 0.5, 1.0, 2.0, or 6.0hr are shown. WT RPKM expression, RPKM expression in *Nfkb1*$^{-/-}$, RPKM expression in *Nfkbiz*$^{-/-}$, or RPKM expression of WT pre-treated with CHX are shown in gray, yellow, green, or red, respectively.

Figure 3-4. RelA, p50, and IκBζ ChIP-seq Analysis

(A) The number of reproducible binding sites (PS > 19 and RPKM > 3 in 2/2 replicates) at each time point is plotted. The total number of unique binding sites for each protein is shown (right).

(B) A pair-wise comparison of the relative binding strength (avg. 2h RPKM) between all p50 and RelA; p50 and IκBζ; or RelA and IκBζ binding sites is shown.

(C) All RelA, p50, and IκBζ peaks are grouped into six kinetic clusters using k-means cluster analysis. The percent of peaks in each cluster is graphed below.

(D) The percent of peaks with open chromatin at 0.0h (PS > 10), closed chromatin at 0.0h (PS < 10), or no ATAC-seq peak is graphed for each kinetic cluster for p50-bound ChIP-seq peaks (top) and IκBζ–bound ChIP-seq peaks in cluster 6 (bottom).

Figure 3-5. Defining Preferential NF-κB Binding and Assessing Transcriptional Relevance

(A) All reproducible 1.0h lipid A stimulated p50 and RelA binding sites (PS > 19 and RPKM > 3 in 2/3 replicates) are ordered based on the ratio between the avg RPKM values of p50 and RelA. The top 2% of binding sites with a preference for p50 or RelA are indicated.

(B) All p50 and RelA binding sites are binned based on p50 versus RelA relative binding strength into 15 bins each containing 290 peaks. In each bin, the percent of peaks that are co-bound by IκBζ (PS > 19 and RPKM > 3 in 2/3 samples) is plotted.

(C) De novo motif analysis was performed for the top 300 binding sites with either p50 or RelA preference. For comparison, de novo motif analysis was also performed on 300 peaks that do not exhibit preferential binding.

(D) All p50 or cRel ChIP-seq (PS > 19 and RPKM > 3 in 2/3 replicates) avg RPKMs are graphed in a scatter plot. All p50-preferential sites when compared to RelA are highlighted in blue.

(E) All p50 or RelA peaks that are within 5 kbps of a TSS are binned based on p50 versus RelA binding strength (avg. 1h RPKM). In each bin, the percent of peaks that annotate to inducible genes (RPKM expression > 3; Induction > 5) and p50-dependent genes (< 33% relative to WT; p-value < 0.01) are plotted.

(F) The IGV genome browser track for the *H2-Aa* promoter reveals a p50-preferential binding site. A time course from 0h – 2.0hr with lipid A stimulation is shown for p50, RelA, and cRel.

Figure 3-6. Investigating the Dependence of IκBζ Binding on p50

(A) For all IκBζ binding sites (PS > 19 and RPKM > 3 in 2/3 replicates) in WT macrophages stimulated with lipid A for 2.0h, the average RPKM in WT versus *Nfkb1*$^{-/-}$ macrophages is plotted. Blue highlights binding sites with > 100% IκBζ RPKM in *Nfkb1*$^{-/-}$ versus WT macrophages. Green highlights IκBζ binding sites with a mild dependence on *Nfkb1* (33% - 100% relative to WT). Orange highlights IκBζ binding sites with a strong dependence on p50 (< 33% binding strength relative to WT). Dots colored red indicate binding sites that annotate to *Nfkb1*/*Nfkbiz*-co-dependent genes.

(B) Quantification of the results in panel A is shown. All IκBζ binding sites are binned based on relative dependence on p50. In each bin, the total number of peaks that annotate to *Nfkb1*-dependent or *Nfkb1*/*Nfkbiz*-co-dependent genes is plotted.

(C) A pie chart reveals the percentage of *Nfkb1*/*Nfkbiz*-co-dependent genes that have a p50-dependent IκBζ binding site within 30 kbps of the TSS. Dark orange shows the total number of co-dependent genes that have a p50-dependent IκBζ binding (< 33% relative to WT) within 30kbps of the TSS. Light orange shows the total number of co-dependent genes that have a p50-semi-dependent IκBζ binding (33% - 66% relative to WT) within 30kbps of the TSS. Brown shows the total number of co-dependent genes that have a p50-independent IκBζ binding (> 66% relative to WT) within 30kbps of the TSS.

(D) Known motif analysis for IκBζ binding sites are binned based on p50-dependent binding. Three unique NF-κB motifs are used in this analysis. The heatmap corresponds to the $-\log_{10}$(p-value) for motif enrichment.

(E) A representative IκBζ binding site with dependence on p50 is displayed. The binding of p50, RelA, and IκBζ as well as H3K27ac is shown at an *Il6* enhancer region for a time course of lipid A stimulation (0h – 2h). The percent maximum peak score is plotted (left).

Figure 3-7. In Vivo and In Vitro Methods to Assess the NF-κB Binding Partner for IκBζ

(A) RelA ChIP-seq was performed in *Nfkb1*$^{-/-}$ macrophages stimulated with lipid A for 1.0h. The pie chart focuses on the 187 IκBζ binding sites with p50-dependence (< 33% relative to WT). Blue highlights peaks where both IκBζ and RelA show strong dependence on p50 (< 33% relative to WT). Light blue highlights p50-dependent IκBζ binding sites where RelA binding is mildly dependent on p50 (33% - 66% relative to WT). Gray highlights p50-dependent IκBζ binding sites where RelA binding is independent of p50 (> 66% relative to WT).

(B) A violin plot shows the average binding strength for RelA Sequential ChIP-seq with a secondary immunoprecipitation for p50, or IκBζ, or as a negative control, p53.

(C) The IGV genome browser track for the *Batf* intronic peaks is a representative example of RelA Sequential ChIP-seq, IκBζ ChIP-seq in WT and *Nfkb1*$^{-/-}$, and RelA ChIP-seq in WT and *Nfkb1*$^{-/-}$ macrophages. The intronic peak with strong p50-dependent IκBζ binding (< 33% binding relative to WT) and p50-dependent RelA binding (< 33% binding relative to WT) is highlighted in gray.

(D) A western blot for the co-immunoprecipitation of RelA with either p50 or IκBζ is shown. Serial dilutions of the input lysate are shown and allow for quantitative analysis. After either no stimulation or 2.0hr stimulation with lipid A, nuclear lysate from the J2-transformed BMDMs was harvested. The bar graph quantifies the percentage of each protein that co-immunoprecipitated with RelA.

Figure 3-8. RNA-seq Analysis from TNF versus Lipid A Stimulated Macrophages

(A) IκBζ expression levels are shown in mRNA (left) or nascent chromatin-associated (right) RNA-seq data sets. BMDMs were stimulated with lipid A (top) or TNF (middle) for 0, 0.5, 1.0, 2.0, or 6.0hr. The bottom row shows the fold change between lipid A stimulation versus TNF stimulation.

(B) The fold change of all NF-κB and IκB members in lipid A-stimulated macrophages relative to TNF-stimulated macrophages is shown. IκBζ is highlighted in blue.

(C) All 132 strongly induced (FC > 10) primary response genes are plotted based on maximum expression in lipid A-stimulated macrophages (x-axis) versus the fold-change in expression

between lipid A-stimulated and TNF-stimulated macrophages (y-axis). Genes with a strong dependence on *Nfkbiz* (< 33% relative to WT), or strong dependence on *Irf3* (< 33% relative to WT) are highlighted in orange and blue, respectively.

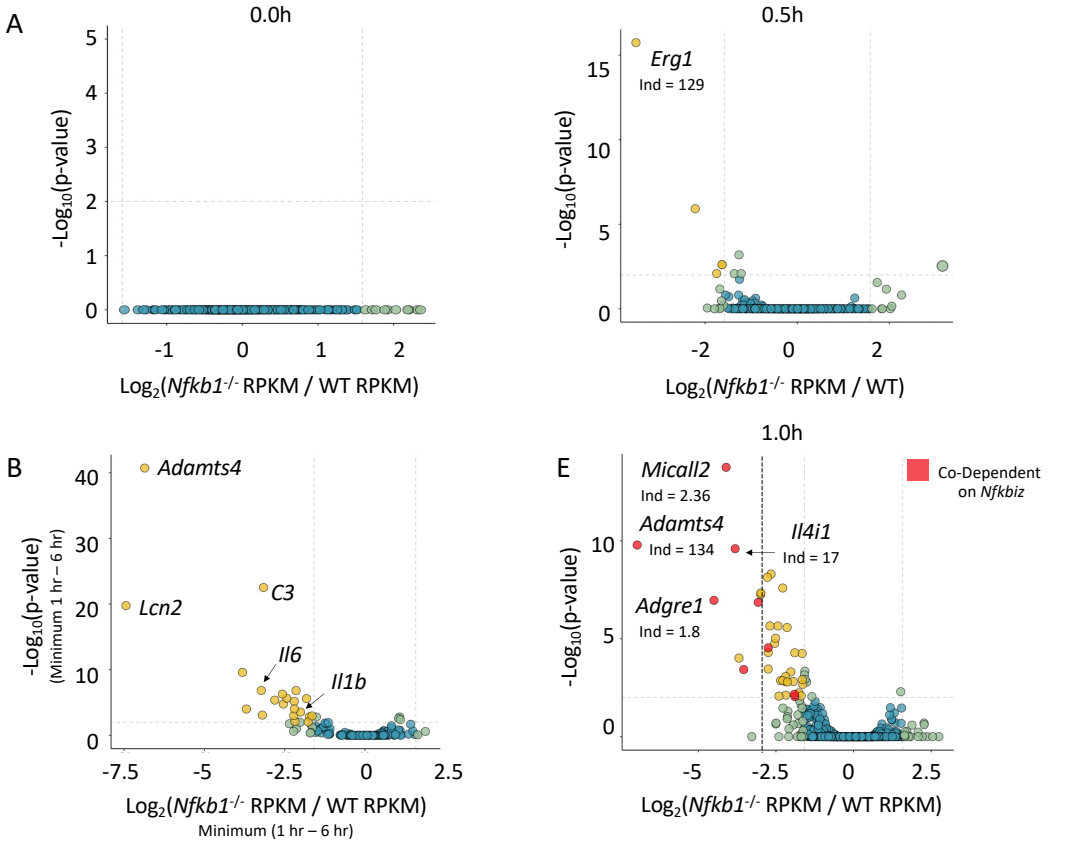
(D) A scatterplot shows all 802 expressed and inducible genes (WT RPKM > 3; Induction > 3-fold) comparing the differential expression between TNF-stimulated and lipid A-stimulated BMDMs (x-axis) versus the percent expression in *Nfkbiz*$^{-/-}$ macrophages relative to WT (y-axis).

(E) Retroviral overexpression of IκBζ was performed in BMDMs. All 802 expressed and inducible genes are graphed based on differential expression in TNF-stimulated and lipid A-stimulated macrophages (x-axis) versus expression in the IκBζ overexpression, TNF-treated BMDMs (y-axis). Genes with a strong dependence on *Nfkbiz* (< 33% relative to WT) are highlighted in red.

Supplemental Figure 3-1: RNA-seq Analysis of ERK-Inhibitor Treated BMDMs

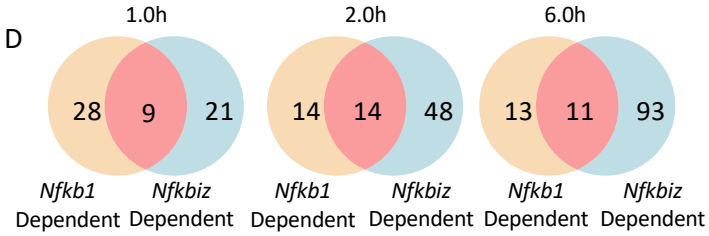
(A) Nascent chromatin-associated RNA-seq was prepared from BMDMs pretreated with an ERK-inhibitor for 1hr prior to stimulation with lipid A. Scatter plots show the relative expression in ERK-inhibitor treated BMDMs versus DMSO-treated BMDMs (x-axis) compared to the induction of expression in DMSO-treated BMDMs treated with lipid A (y-axis).

Figure 3-1. Nascent RNA-seq Analysis of *Nfkb1*$^{-/-}$, *Nfkbiz*$^{-/-}$, *Bcl3*$^{-/-}$, and *Nfkbid*$^{-/-}$ BMDMs



C

Genotype	0.0h	0.5h	1.0h	2.0h	6.0h
Nfkb1$^{-/-}$	-	4	37	28	24
Nfkbiz$^{-/-}$	-	-	30	62	104
Bcl3$^{-/-}$	-	-	-	-	-
Nfkbid$^{-/-}$	-	-	-	-	-



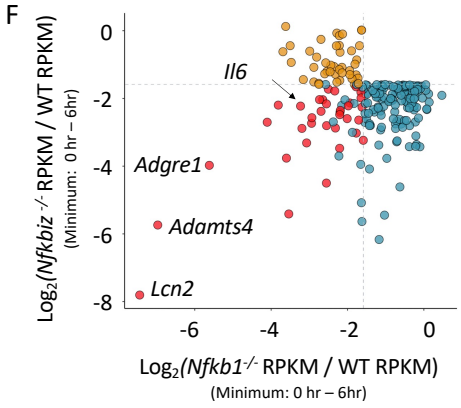
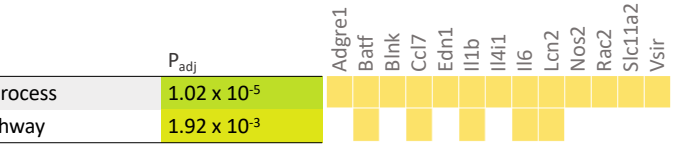


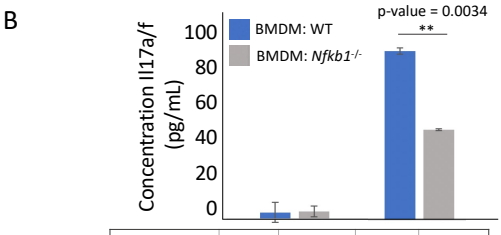
Figure 3-2. Gene Ontology and Biological Significance of *Nfkb1*-Dependent Genes

A

ID	Source	Term Name	P_{adj}
1	GO:BP	Immune System Process	1.02×10^{-5}
2	KEGG	Il-17 Signaling Pathway	1.92×10^{-3}



B



BMDM Conditions				
Genotype:	WT	*Nfkb1*$^{-/-}$	WT	*Nfkb1*$^{-/-}$
Stimulation:	Unstimulated		LPS	

T-Cell Conditions		
T-Cell	WT	WT
Anti-CD3 (plate-bound)	10ug/ml	10ug/mL
Anti-CD28	2ug/mL	2ug/mL
BMDM Supernatant	100uL	100uL

Figure 3-3. Kinetic Analysis of Nascent RNA-seq from WT, *Nfkb1*$^{-/-}$, or *Nfkbiz*$^{-/-}$ BMDMs

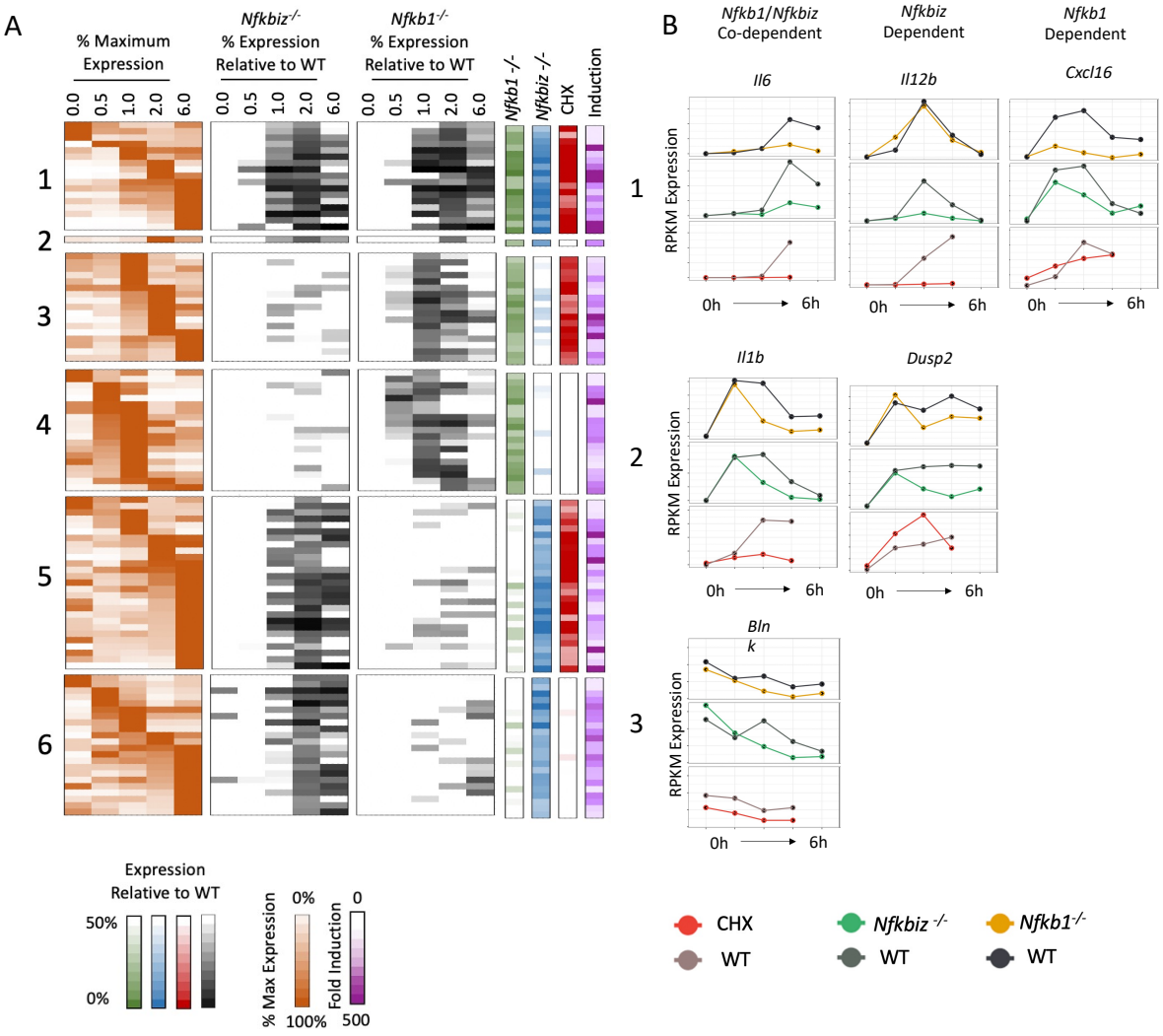


Figure 3-4. RelA, p50, and IκBζ ChIP-seq Analysis

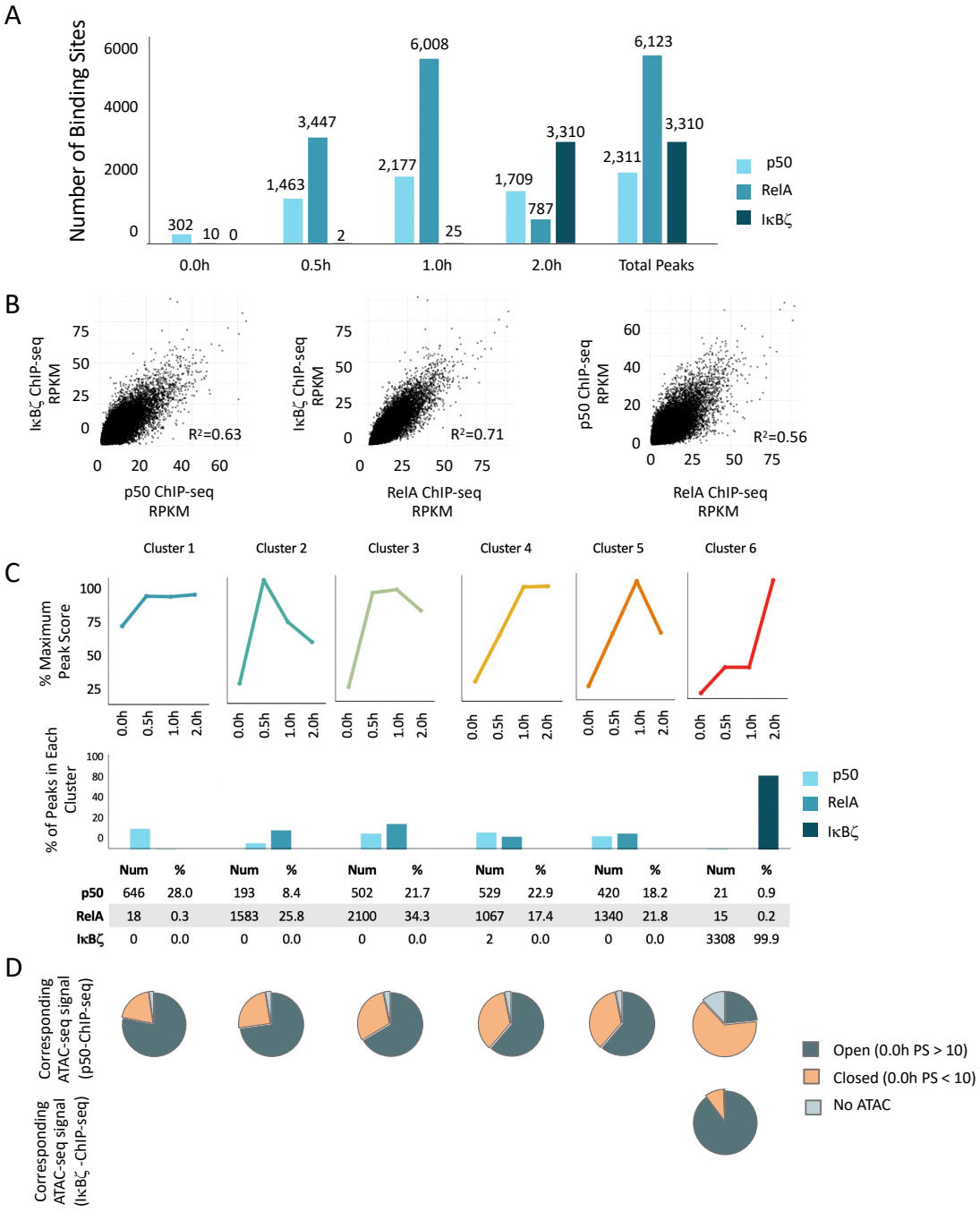


Figure 3-5. Defining Preferential NF-κB Binding and Assessing Transcriptional Relevance

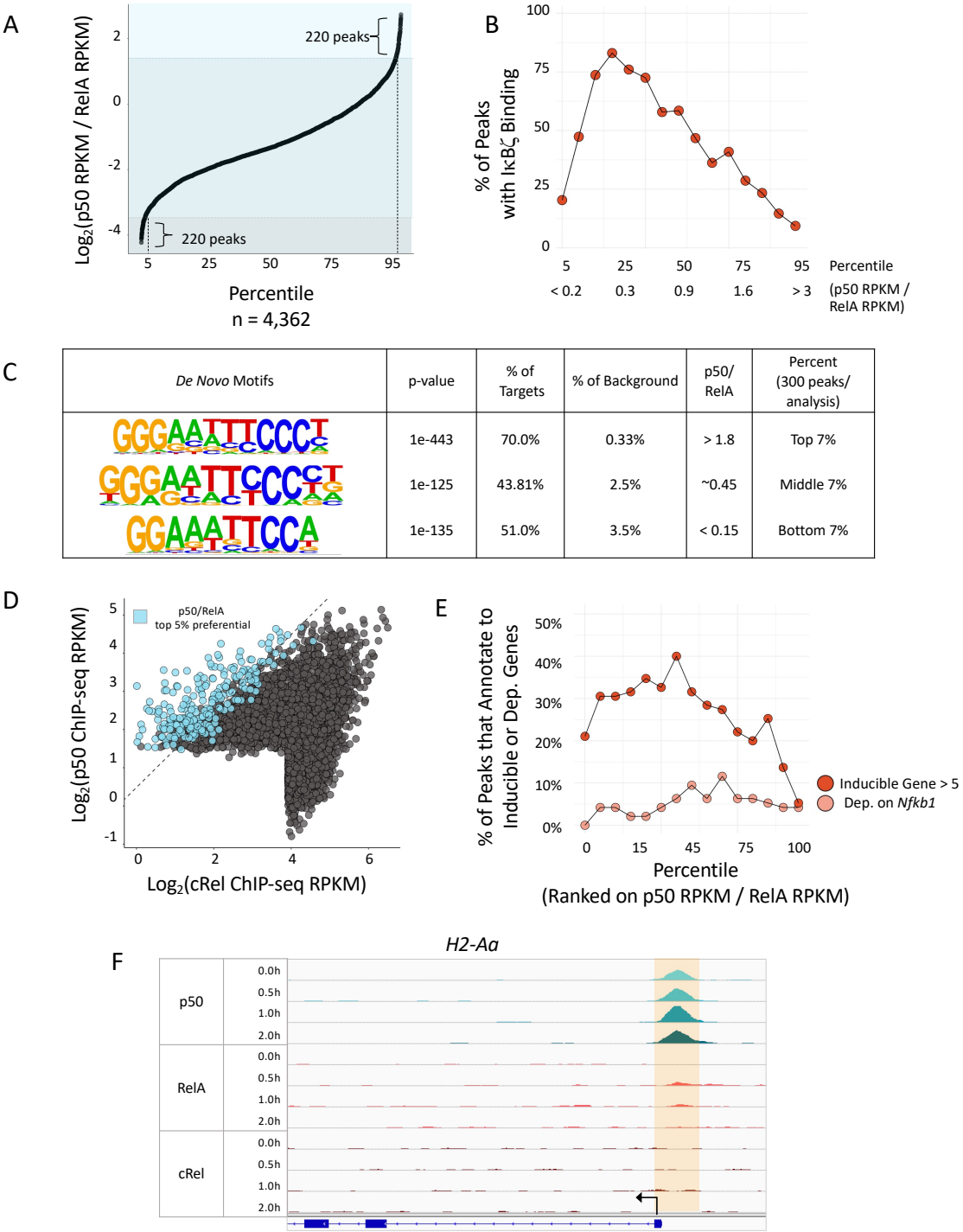


Figure 3-6. Investigating the Dependence of IκBζ Binding on p50

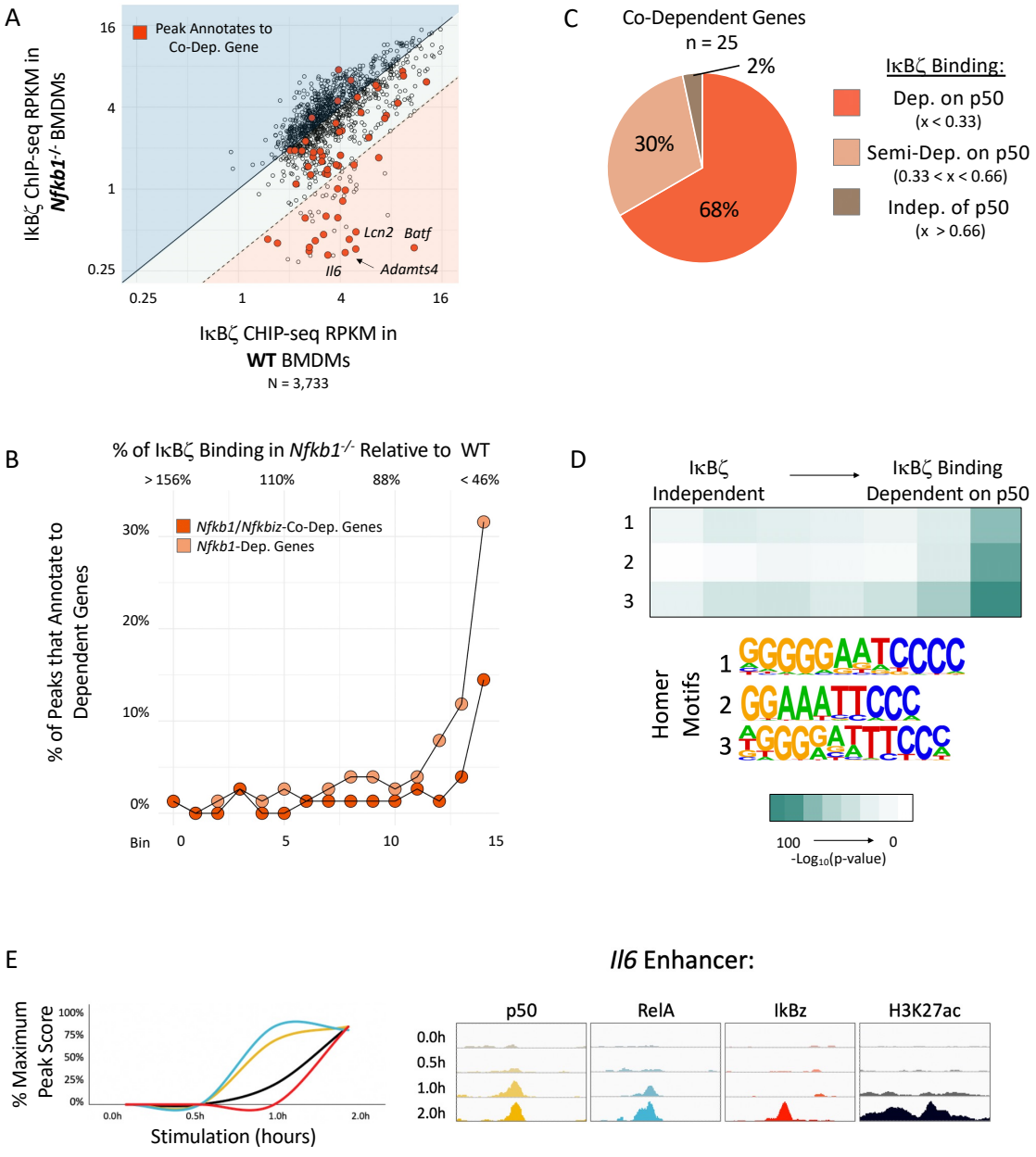


Figure 3-7. In Vivo and In Vitro Methods to Assess the NF-κB Binding Partner for IκBζ

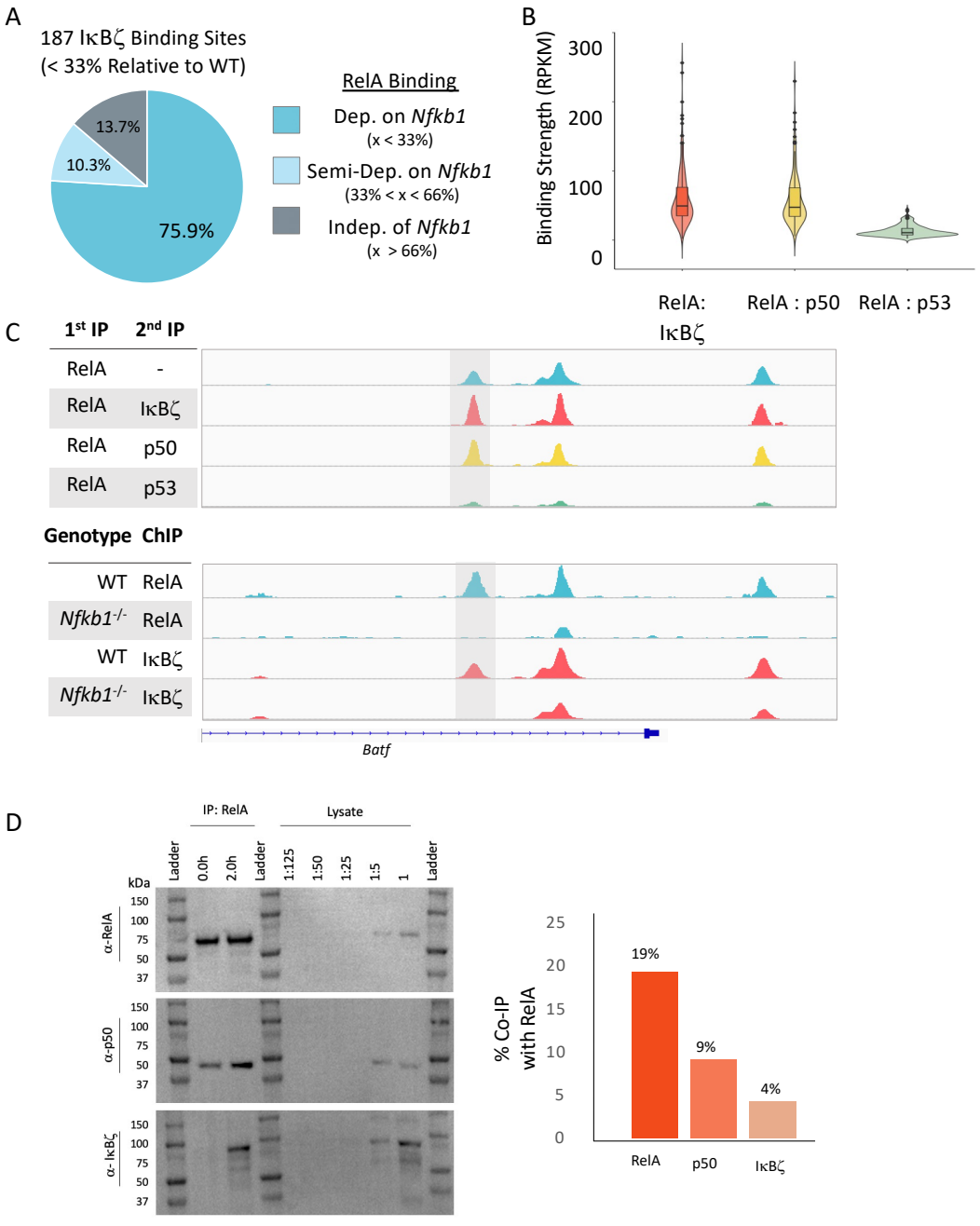
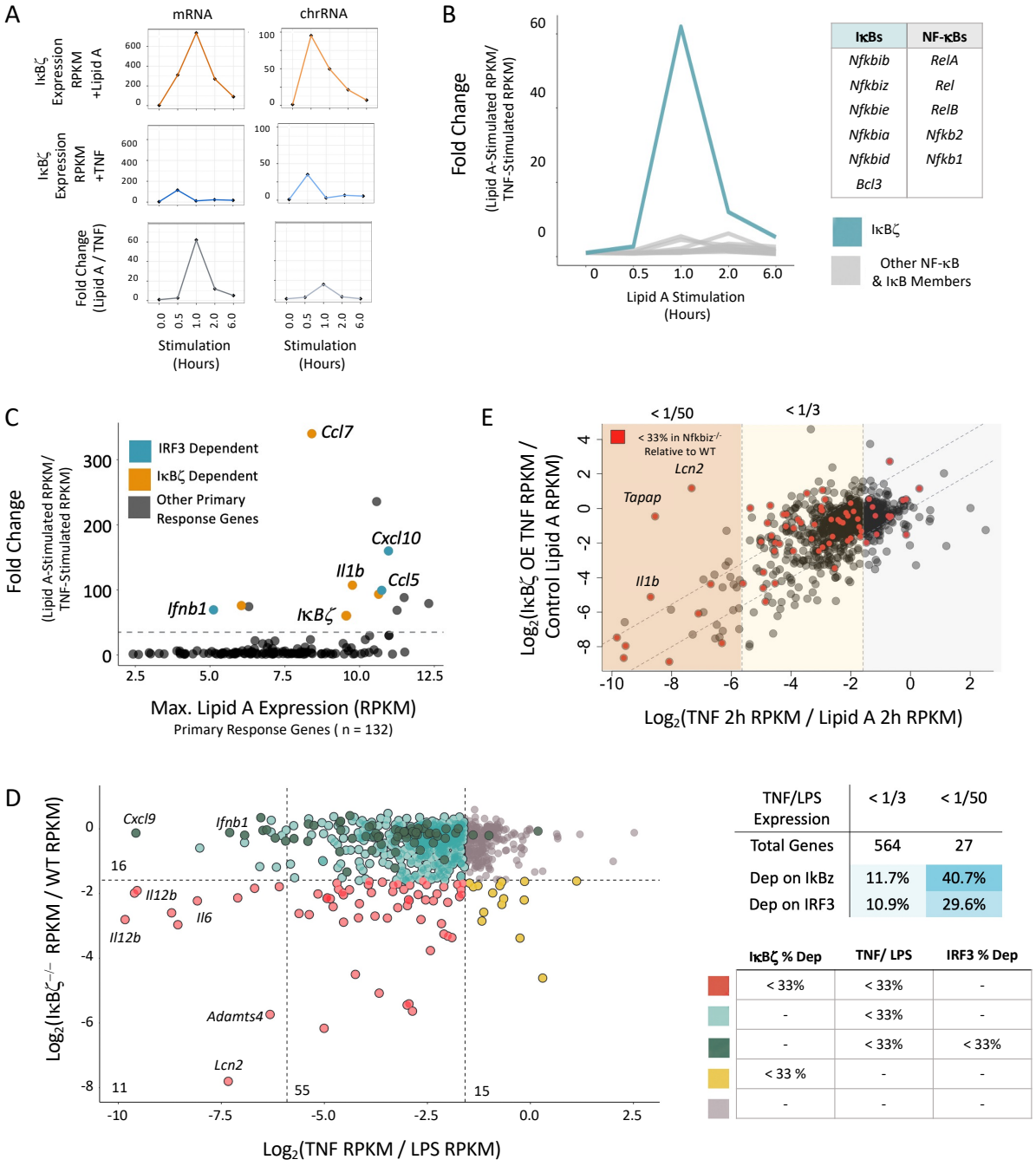
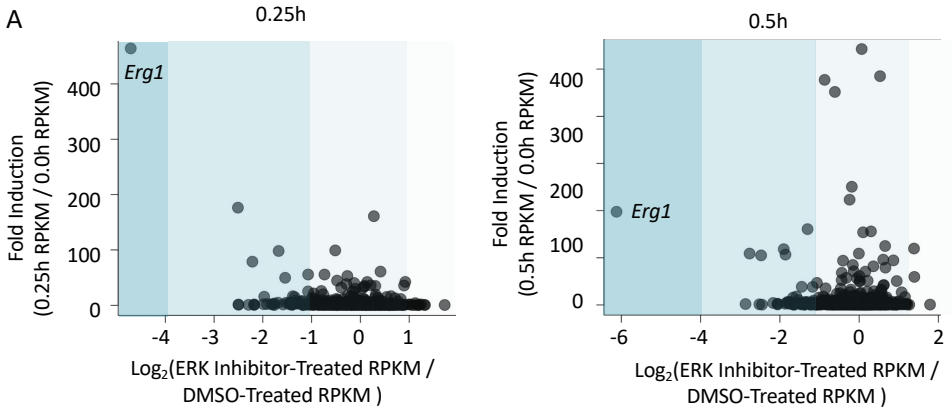


Figure 3-8. RNA-seq Analysis from TNF versus Lipid A Stimulated Macrophages



Supplemental Figure 3-1. RNA-seq Analysis of ERK-Inhibitor Treated BMDMs



WORKS CITED

1. Janeway, C.A., ravers, P, Walport, M, et al. The components of the immune system. Immunobiol. Immune Syst. Heal. Dis. 5th Ed. New York Garl. Sci. (2001).
2. Mass, E., Nimmerjahn, F., Kierdorf, K. & Schlitzer, A. Tissue-specific macrophages: how they develop and choreograph tissue biology. Nat Rev Immunol (2023). doi:10.1038/s41577-023-00848-y
3. Werner, S. L., Barken, D. & Hoffmann, A. Stimulus specificity of gene expression programs determined by temporal control of IKK activity. Science (80-.). 309, 1857–1861 (2005).
4. Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. Cell 140, 805–820 (2010).
5. Jost, P. J. & Ruland, J. Aberrant NF-κB signaling in lymphoma: mechanisms, consequences, and therapeutic implications. Blood 109, 2700–2707 (2007).
6. O'Neill, L. A. J., Golenbock, D. & Bowie, A. G. The history of Toll-like receptors-redefining innate immunity. Nat. Rev. Immunol. 13, 453–460 (2013).
7. Crabtree, G. R. & Hargreaves, D. C. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Nat. Publ. Gr. 21, 396–420 (2011).
8. Brenner, D., Blaser, H. & Mak, T. W. Regulation of tumour necrosis factor signalling: live or let die. Nat. Rev. Immunol. 2015 156 15, 362–374 (2015).
9. Aggarwal, B. B. et al. Human tumor necrosis factor. Production, purification, and characterization. J. Biol. Chem. 260, 2345–2354 (1985).
10. Van Loo, G. & Bertrand, M. J. M. Death by TNF: a road to inflammation. Nat Rev Immunol (2022). doi:10.1038/s41577-022-00792-3

11. Sen, S., Cheng, Z., Sheu, K. M., Chen, Y. H. & Hoffmann, A. Gene regulatory strategies that decode the duration of NF-kappa B dynamics contribute to LPS- versus TNF-specific gene expression. Cell Syst. (2019). doi:10.1016/j.cels.2019.12.004

12. Yamamoto, M. et al. A novel Toll/IL-1 receptor domain-containing adapter that preferentially activates the IFN-beta promoter in the Toll-like receptor signaling. J. Immunol. 169, 6668–6672 (2002).

13. Hayden, M. S. & Ghosh, S. Signaling to NF-kappaB. Genes Dev. 18, 2195–2224 (2004).

14. Oeckinghaus, A. & Ghosh, S. The NF-kappaB family of transcription factors and its regulation. Cold Spring Harbor Perspectives in Biology 1, (2009).

15. Liu, T., Zhang, L., Joo, D. & Sun, S. C. NF-κB signaling in inflammation. Signal Transduct. Target. Ther. 2, 17023 (2017).

16. O'Dea, E. & Hoffmann, A. The regulatory logic of the NF-kappaB signaling system. Cold Spring Harb. Perspect. Biol. 2, (2010).

17. Shih, V. F., Tsui, R., Caldwell, A. & Hoffmann, A. A single NF-κB system for both canonical and non-canonical signaling. Cell Res. 21, 86 (2011).

18. Zhang, Q., Lenardo, M. J. & Baltimore, D. 30 years of NF-κB: a blossoming of relevance to human pathobiology. Cell 168, 37–57 (2017).

19. Cheng, C. S. et al. The specificity of innate immune responses is enforced by repression of interferon response elements by NF-κB p50. Sci. Signal. 4, ra11 (2011).

20. Zhao, X., Ross, E. J., Wang, Y. & Horwitz, B. H. Nfkb1 inhibits LPS-induced IFN-β and IL-12 p40 production in macrophages by distinct mechanisms. PLoS One 7, e32811 (2012).

21. Elsharkawy, A. M. et al. The NF-kappaB p50:p50:HDAC-1 repressor complex orchestrates transcriptional inhibition of multiple pro-inflammatory genes. J. Hepatol. 53, 519–27 (2010).

22. Cao, S., Zhang, X., Edwards, J. P. & Mosser, D. M. NF-kappaB1 (p50) homodimers differentially regulate pro- and anti-inflammatory cytokines in macrophages. J. Biol. Chem. 281, 26041–50 (2006).

23. Bours, V. et al. The oncoprotein Bcl-3 directly transactivates through kapp B motifs via association with DNA-binding p50B homodimers. Cell 72, 729–739 (1993).

24. Fujita, T., Nolan, G. P., Liou, H.-C., Scott, M. L. & Baltimore, D. The candidate proto-oncogene bcl-3 encodes a transcriptional coactivator that activates through NF-kappa B p50 homodimers.

25. Motoyama, M., Yamazaki, S., Eto-Kimura, A., Takeshige, K. & Muta, T. Positive and negative regulation of nuclear factor-κB-mediated transcription by IκB-ζ, an inducible nuclear protein. J. Biol. Chem. 280, 7444–7451 (2005).

26. Yamamoto, M. et al. Regulation of Toll/IL-1-receptor mediated gene expression by the inducible nuclear protein IκBzeta. Nature 430, (2004).

27. Trinh, D. V., Zhu, N., Farhang, G., Kim, B. J. & Huxford, T. The nuclear IκB protein IκBζ specifically binds NF-κB p50 homodimers and forms a ternary complex on κB DNA. J. Mol. Biol. 379, 122–135 (2008).

28. Hayden, M. S. & Ghosh, S. NF-κB, the first quarter-century: remarkable progress and outstanding questions. Genes Dev. 26, 203–234 (2012).

29. Motoyama, M., Yamazaki, S., Eto-Kimura, A., Takeshige, K. & Muta, T. Positive and negative regulation of nuclear factor-kappaB-mediated transcription by IkappaB-zeta, an inducible nuclear protein. J. Biol. Chem. 280, 7444–51 (2005).

30. Yamamoto, M. et al. Regulation of Toll/IL-1-receptor-mediated gene expression by the inducible nuclear protein IκBζ. Nature 430, 218–222 (2004).

31. Yamazaki, S., Muta, T., Matsuo, S. & Takeshige, K. Stimulus-specific induction of a novel nuclear factor-kappaB regulator, IkappaB-zeta, via Toll/Interleukin-1 receptor is mediated by mRNA stabilization. J. Biol. Chem. 280, 1678–1687 (2005).

32. Eto, A., Muta, T., Yamazaki, S. & Takeshige, K. Essential roles for NF-κB and a Toll/IL-1 receptor domain-specific signal(s) in the induction of IκB-ζ. Biochem. Biophys. Res. Commun. 301, 495–501 (2003).

33. Fitzgerald, K. A. et al. LPS-TLR4 signaling to IRF-3/7 and NF-kappaB involves the toll adapters TRAM and TRIF. J. Exp. Med. 198, 1043–1055 (2003).

34. Gantke, T., Sriskantharajah, S. & Ley, S. C. Regulation and function of TPL-2, an IκB kinase-regulated MAP kinase kinase kinase. Cell Res. 2011 211 21, 131–145 (2010).

35. Waterfield, M. R., Zhang, M., Norman, L. P. & Sun, S. C. NF-κB1/p105 regulates lipopolysaccharide-stimulated MAP kinase signaling by governing the stability and function of the Tpl2kinase. Mol. Cell 11, 685–694 (2003).

36. Blair, L. et al. TPL-2 inhibits IFN-β expression via an ERK1/2-TCF-FOS axis in TLR4-stimulated macrophages. J. Immunol. 208, 941–954 (2022).

37. Tong, A. J., Liu, X. et al. A stringent systems approach uncovers gene-specific mechanisms regulating inflammation. Cell 165, 165–179 (2016).

38. Kannan, Y. et al. IκBζ augments IL-12- and IL-18-mediated IFN-γ production in human NK cells. Blood 117, 2855–2863 (2011).

39. Yamazaki, S., Muta, T. & Takeshige, K. A novel IκB protein, IκB-ζ, induced by proinflammatory stimuli, negatively regulates nuclear factor-κB in the nuclei. J. Biol. Chem. 276, 27657–27662 (2001).

40. Grondona, P. et al. Threonine phosphorylation of IκBζ mediates inhibition of selective proinflammatory target genes. J. Invest. Dermatol. 140, 1805-1814.e6 (2020).

41. Purbey, P. K. et al. Defined sensing mechanisms and signaling pathways contribute to the global inflammatory gene expression output elicited by ionizing radiation. Immunity 47, 421-434.e3 (2017).

CHAPTER 4

Stepwise Neofunctionalization of c-Rel during Vertebrate Evolution

Allison E. Daly,[1,2,10] Abraham B. Chang,[1,2,10] Prabhat Purbey,[1-3,10] Kevin J. Williams,[1,2,10] George Yeh,[1,2] Shuxing Li,[6] Yongquing Wu,[6] Scott D. Pope,[1,2] Byrappa Venkatesh,[7] Benjamin D. Redelings,[8] Sibon Li,[4] Kaylin Nguyen,[1] Joseph Rodrigues,[1] Marc A. Suchard,[4] Trevor Siggers,[9] Lin Chen,[6] and Stephen T. Smale[1-3,5*]

[1]Department of Microbiology, Immunology, and Molecular Genetics
[2]Molecular Biology Institute
[3]Department of Medicine
[4]Department of Human Genetics
[5]Howard Hughes Medical Institute
University of California, Los Angeles, CA 90095, USA
[6]Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089
[7]Comparative Genomics Lab, Institute of Molecular and Cell Biology, Singapore 138673
[8]Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045
[9]Department of Biology, Boston University, Boston, MA 02215
[*]Correspondence: *smale@mednet.ucla.edu* (Tel: 310-206-4777; Fax: 310-206-8623)

[10]These authors contributed equally.

SUMMARY

Adaptive immunity and the five vertebrate NF-κB/Rel family members first appeared in cartilaginous fish, suggesting that NF-κB family expansion allowed the acquisition of new functions to regulate adaptive immune responses. Transcriptome profiling revealed that, even in macrophages, the NF-κB family member, c-Rel, most potently regulates a cytokine gene linked to adaptive immunity, *Il12b*, with limiting roles at key regulators of innate immunity. Neofunctionalization of c-Rel to regulate *Il12b* depends on its unique DNA-binding properties, which we examined using structural, biochemical, functional, and genomic approaches. Among our findings was functional c-Rel homodimer binding to motifs with little resemblance to canonical NF-κB motifs. To determine whether c-Rel's unique binding properties drove c-Rel-RelA divergence, we compared binding properties in various vertebrate species. c-Rel-RelA binding properties diverged in mammals and amphibians but were comparable in earlier vertebrates, suggesting that divergent DNA binding emerged relatively late during vertebrate evolution to support the increasing complexity of adaptive immune regulation.

INTRODUCTION

The complexity of the vertebrate adaptive immune system reflects billions of years of evolution. Properly controlled adaptive immune responses are now known to rely on complex interactions between dozens or hundreds of cell types. The V(D)J adaptive immune system, first observed in cartilaginous fish that emerged approximately 450 million years ago, has been retained in all jawed vertebrates (Gnathostomata) that have been studied, with more primitive jawless fish (Cyclostomata), including lampreys and hagfish, possessing a distinct mechanism to support adaptive immune responses (Carmona and Schatz, 2017; Boehm et al., 2018).

Although the V(D)J adaptive immune system is fundamentally similar among Gnathostomata, differences have been reported and may be essential to accommodate the unique needs of distinct vertebrate groups, such as fetal tolerance in placental mammals (Rackaityte and Halkias, 2020). Consistent with these unique needs, a placental mammal-specific enhancer that supports extrathymic generation of T-regulatory (Treg) cells has been identified in the locus encoding FoxP3, a key transcription factor controlling Treg development (Samstein et al., 2012). Functionally impactful amino acid changes in the FoxP3 protein were also acquired during distinct stages of vertebrate evolution (Andersen et al., 2012).

Multiple mechanisms by which gene duplication and specialization can support evolution have been described (Conant and Wolfe, 2008; Andersson et al., 2015). Subfunctionalization occurs when the functions of an ancestral protein are divided between its gene-duplicated descendants. Neofunctionalization represents the emergence of new functions in one or more descendants. Neofunctionalization often arises as a result of changes in the expression patterns of the duplicated genes, but mutations resulting in amino acid changes also allow new functions to emerge (Carroll, 2005; Lynch and Wagner, 2008; Cheatle Jarvela and Hinman, 2015). A comprehensive study of transcription factor paralogs in yeast suggested that these amino acid changes usually occur outside the DNA-binding domains of paralog pairs (Gera et al., 2022).

The vertebrate NF-κB/Rel transcription factor family is comprised of five members that play prominent roles in the regulation of both innate and adaptive immune responses (Sen and

Baltimore, 1986; Hayden and Ghosh, 2012). The smaller number of NF-κB ancestors found in invertebrates possess well-documented roles in development, homeostasis, and innate immunity (Williams and Gilmore, 2020). NF-κB proteins are characterized by an N-terminal Rel homology region of approximately 300 amino acids that supports sequence-specific DNA binding and assembly into a diverse array of homodimeric and heterodimeric species (Hayden and Ghosh, 2012). Most dimers are retained in the cytoplasmic of resting cells in association with IκB inhibitor proteins, with IκB degradation and the nuclear translocation of the NF-κB dimers directed by signaling pathways upon sensing of microbial and environmental threats. Loss-of-function studies have revealed distinct biological functions for each family member (Gerondakis et al. 1999, Gerondakis and Siebenlist, 2010; Hayden and Ghosh, 2012), but much less is known about the target genes and regulatory mechanisms for each dimeric species.

RelA and c-Rel are the two most closely related NF-κB family members, with highly conserved RHRs and identical DNA-contacting residues, but divergent C-terminal activation domains. RelA is abundant in most mammalian cell types, whereas high c-Rel expression is found primarily in hematopoietic lineages (Gilmore and Gerondakis, 2011). Mice deficient in the *Rela* gene (encoding RelA) exhibit embryonic lethality due to broad deficiencies in cell survival (Hayden and Ghosh, 2012). In contrast, mice deficient in the *Rel* gene (encoding c-Rel) exhibit a variety of immune abnormalities, including diminished T and B-cell responses, defective T regulatory cell (Treg) development, and a prominent loss of both Th1 and Th17 immune responses (Köntgen et al., 1995; Tumong et al., 1998; Gilmore and Gerondakis, 2011; Hayden and Ghosh, 2012).

In immune cells that express both RelA and c-Rel, the two proteins often act redundantly to regulate inducible genes (Hoffmann et al., 2006; Hayden and Ghosh, 2012). Nevertheless, c-Rel has been reported to contribute in a non-redundant manner to the induction of several genes (Gilmore and Gerondakis, 2011). Although quantitative genome-wide studies of c-Rel's relative importance at different proposed target genes have not been reported. We and others previously described a potent role for c-Rel in the activation of the *Il12b* gene in both mouse and human antigen-presenting cells, including macrophages and some dendritic cell subsets (Sanjabi et al.,

2000; Mason et al., 2002; Wang et al., 2007; Lévy et al., 2021). In a chimeric protein analysis, the c-Rel and RelA activation domains were interchangeable (Sanjabi et al., 2005). Instead, the c-Rel requirement for *Il12b* induction localized to an 86-residue portion of the RHR containing 46 amino acid differences between RelA and c-Rel (Sanjabi et al., 2005). Despite identical DNA-contacting residues in RelA and c-Rel (Chen and Ghosh, 1999; Huang et al., 2001), these 46 residues allowed c-Rel homodimers to bind consensus and non-consensus NF-κB motifs with a higher affinity than RelA homodimers (Sanjabi et al., 2005; Siggers et al., 2011). The finding that the same small region of c-Rel required for *Il12b* gene induction in *Rel*$^{-/-}$ macrophages also confers a large DNA affinity difference suggests that the affinity difference is responsible for the functional difference.

In this study, we extended our knowledge of the unique properties of c-Rel and the relevance of these properties to the functional specialization of RelA and c-Rel during vertebrate evolution. We first found that *Il12b* exhibits an unusually strong dependence on c-Rel in activated macrophages, with only a small number of other inducible genes exhibiting c-Rel dependence. A protein-DNA co-crystal structure of a c-Rel-RelA chimera provided insights into the structural basis of c-Rel's unique binding properties. We then found by protein-binding microarray (PBM) and surface plasmon resonance (SPR) analyses that c-Rel homodimers can bind motifs that bear little resemblance to canonical NF-κB motifs and previously would have been unrecognizable. *Il12b* transcriptional induction requires a tandem array of divergent motifs in the *Il12b* promoter that bind c-Rel homodimers with a high degree of selectivity and cooperativity. Despite the clear differences between the intrinsic DNA-binding properties of mouse RelA and c-Rel, and despite the strong evidence that the affinity difference is functionally important, the affinity difference was observed only with RelA and c-Rel DNA-binding domains from mice, humans, and frogs, with no significant affinity difference observed with the DNA-binding domains from elephant shark, zebrafish, and chicken. These results suggest that this critical biochemical difference was not responsible for the initial functional specialization of RelA and c-Rel in early vertebrates, but it

instead emerged at a later stage of vertebrate evolution to support further specialization of the adaptive immune system.

RESULTS

Evolution of the NF-κB Family

A full understanding of the NF-κB family of transcription factors will require an understanding of the family's evolution. As an initial step toward this goal, we prepared a phylogenetic tree based on predicted amino acid sequences of the RHRs of NF-κB family members from several eukaryotic species (Figure 4-1A). This analysis revealed that the five NF-κB family members found in mammals and other vertebrate species - p50 (*Nfkb1*), p52 (*Nfkb2*), RelA (*Rela*), c-Rel (*Rel*), and RelB (*Relb*) - are also conserved in elephant shark (Venkatesh et al., 2014), representative of cartilaginous fish (Chondrichthyes), the most primitive gnathostome class. Notably, the elephant shark genome encodes a sixth NF-κB protein that exhibits the closest homology to its p50 and p52 paralogs (Figure 4-1A). Ancestral invertebrate NF-κB proteins (typically two or three in each species) diverge considerably from the vertebrate RelA, c-Rel, and RelB clusters (Figure 4-1A). A closer phylogenetic relationship is apparent between the vertebrate p50 and p52 proteins and an NF-κB protein found in at least two invertebrate species, sea squirts, and sea urchins (Figure 4-1A); these invertebrate and vertebrate proteins may therefore be derived from a common invertebrate ancestor.

Because RelA, c-Rel, and RelB are evolutionarily distant from the invertebrate NF-κBs, it was of interest to examine NF-κB proteins in lamprey, which represent Cyclostomata (jawless fish), which likely diverged from jawed, cartilaginous fish during the Ordovician period, roughly 450 million years ago (Boehm et al., 2018). Lamprey genome sequences are now extensive (Smith et al., 2018), but remain incomplete. We therefore prepared a cDNA library from lamprey blood and successfully isolated and sequenced the genes for five NF-κB family members using PCR and degenerate PCR; a subset of these sequences are apparent in public sea lamprey genome sequences (Smith et al., 2018).

One of the lamprey genes clustered relatively closely with two of the elephant shark p50/p52 proteins (Figure 4-1A). This gene may share a common ancestor with both vertebrate p50 and p52. However, none of the five predicted lamprey NF-κBs clustered closely with the RelA, c-Rel, and RelB proteins from other vertebrates. These results suggest that RelA, c-Rel, and RelB first appeared after the divergence of ancestral jawless and jawed fish. It is difficult to predict the specific evolutionary events that led to the emergence of the *Rela*, *Rel*, and *Relb* genes from the NF-κB genes found in invertebrates and jawless fish.

Extensive evidence suggests that the evolution of jawed fish was accompanied by two whole-genome duplication events (Holland and Ocampo Daza, 2018). This knowledge suggests that the *Rela* and *Rel* genes emerged following the duplication of a common ancestral gene during one of these whole-genome duplications. The *Relb* gene may also share a common ancestral gene with *Rela* and *Rel*. Importantly, because the V(D)J adaptive immune system is first observed in cartilaginous fish, the divergence and neofunctionalization of the five NF-κB family members may have provided support for adaptive immunity.

Highly Selective Role of c-Rel in Lipid A-Stimulated BMDMs

RelA and c-Rel contain the two most closely related RHRs among the five vertebrate NF-κB family members, including identical DNA-contacting residues (Chen and Ghosh, 1999; Huang et al., 2001). Despite considerable redundancy between the two proteins in some settings, they also contribute unique functions for the regulation of adaptive immunity (Wang et al., 2007; Gerondakis and Siebenlist, 2010; Hayden and Ghosh, 2018). To improve our understanding of c-Rel's functions in the regulation of immune responses, we performed RNA-seq with wild-type C57BL/6 and *Rel*$^{-/-}$ bone marrow-derived macrophages (BMDMs) stimulated with lipid A for 0, 60, and 120 min to identify c-Rel-dependent genes at a genome-wide scale. This analysis was done by both nascent transcript (chromatin-associated) RNA-seq (Figure 4-1B), to allow a quantitative analysis of c-Rel's impact on transcription (because mRNA levels can also be strongly influenced by transcript stability), and also by mRNA-seq (data not shown). Surprisingly, these analyses

revealed c-Rel-dependent expression of only a small number of genes one and two hr post-stimulation, when direct targets are expected to first be induced (Figure 4-1B). *Il12b* was found to exhibit unusually strong c-Rel-dependence at both time points, and this gene was induced by lipid A much more potently than the other genes that exhibited strong c-Rel-dependence: these genes include *Clcf1* (encoding cardiotrophin-like cytokine factor 1), *Orai2* (encoding a calcium channel component), *Il4i1* (encoding a secreted L-amino acid oxidase), *Noct* (encoding a member of the exonuclease-endonuclease-phosphatase superfamily of enzymes), and *Tnfsf9* (encoding a transmembrane cytokine of the tumor necrosis factor family) (Figure 4-1B, fold induction values in parenthesis). The functions of these latter genes in cells of the innate immune system remain poorly understood.

The small number of lipid A-induced genes that exhibit strong c-Rel dependence is consistent with the view that RelA and c-Rel act with considerable redundancy in the activation of pro-inflammatory genes in macrophages, but with c-Rel critical for the induction of *Il12b* and a very small number of other genes. The primary function of the IL-12b (IL-12 p40) protein is to regulate adaptive immune responses by promoting the differentiation of naïve T cells into the Th1 and Th17 lineages. Moreover, the regulation of *Il12b* expression by c-Rel has been shown to be critical for the development of Th1 and Th17 cells (Ruan et al., 2011; Zhang et al., 2017; Lévy et al., 2021). Thus, c-Rel's highly selective role as a potent inducer of *Il12b* is consistent with the hypothesis that, during vertebrate evolution, c-Rel's neofunctionalization allowed it, in part, to support adaptive immunity via its critical *Il12b* regulatory function in macrophages, along with its well-documented functions in T and B cell subsets (Gerondakis and Siebenlist, 2010; Gilmore and Gerondakis, 2011).

Structural Analysis of the c-Rel versus RelA DNA Affinity Difference

As described above (see Introduction), we previously found that an 86-residue region of the c-Rel RHR, containing 46 amino acid differences between c-Rel and RelA, is responsible for c-Rel's unique ability to activate the *Il12b* transcription in *Rel*$^{-/-}$ macrophages (Sanjabi et al., 2005). We

also showed that these same residues allow c-Rel homodimers to bind NF-κB motifs with an order of magnitude higher affinity than RelA homodimers (Sanjabi et al., 2005; Siggers et al., 2011). To understand the structural basis for this intrinsic affinity difference, we solved an X-ray co-crystal structure of a DNA-bound RelA-c-Rel chimeric protein, RelA(C46), in which the 46 key residues of c-Rel were introduced into the RelA RHR in place of the corresponding RelA RHR. We previously found that the DNA-binding affinity of this chimeric protein containing only 46 c-Rel RHR residues is comparable to that of c-Rel (Sanjabi et al. 2005).

The RelA(C46) co-crystal structure was solved using the same oligonucleotide sequence that was previously used for the RelA-DNA co-crystal structure, allowing a direct comparison (data not shown). As expected, based on prior RelA and c-Rel structures, the DNA-contacting residues of the RelA and RelA(C46) proteins are identical. However, clear structural differences were observed in the orientation of the DNA-contacting residues relative to critical nucleotide bases (data not shown), suggesting that these differences strongly impact the affinity of the protein-DNA interaction.

Specific Binding of c-Rel Homodimers to Highly Divergent DNA Recognition Motifs

Protein-binding microarrays (PBMs) previously revealed that c-Rel and RelA homodimers bind similar distributions of DNA recognition motifs (Figure 4-2B, black dots; Siggers et al., 2011). However, SPR experiments revealed that c-Rel homodimers bind with a much higher affinity than RelA homodimers to the full range of sequences evaluated by PBM (Siggers et al., 2011). That is, the PBM results only show the relative binding of an individual protein to the range of DNA motifs present in the array, but SPR allows a comparison of the relative affinities of c-Rel homodimers and RelA homodimers for specific oligonucleotide sequences. The higher c-Rel homodimer binding affinity allows c-Rel homodimer binding to motifs that diverge from the NF-κB consensus to be detected more readily than RelA homodimer binding, as observed in early comparative studies (Kunsch et al., 1992).

Interestingly, although most sequences bound by c-Rel and RelA homodimers in the PBM experiments are recognizable as consensus or non-consensus NF-κB motifs, close scrutiny of the prior PBM profiles revealed that many oligonucleotides with five tandem T:A bps were also bound by both c-Rel and RelA homodimers (Figure 4-2B, top right, red dots). Preferential binding was not observed with many other DNA motifs that do not resemble known NF-κB recognition motifs (e.g. GAGAT and CACTA in Figure 4-2B) or a motif that preferentially binds NF-κB p50 homodimers (Figure 4-2B, GGGGG; Siggers et al., 2012).

Frequent binding to oligonucleotides with tandem T:A bps was unexpected because structural and biochemical studies have revealed that NF-κB binding energy is primarily due to contacts with G:C bps in both half-sites of a dimeric recognition sequence (Chen and Ghosh, 1999). Nevertheless, SPR revealed that five T:A bps following an NF-κB half-site with five G:C bps results in a much slower c-Rel homodimer off-rate in comparison to an oligonucleotide in which the five G:C bps were followed by a different sequence (Figure 4-2C). RelA homodimer binding to both oligonucleotides was much weaker by SPR (Figure 4-2C). The binding difference between c-Rel and RelA homodimers was largely due to the 46 key residues of c-Rel described above, as binding of the RelA(C46) chimeric protein was readily detected by SPR, with an off-rate that was only moderately reduced in comparison to wild-type c-Rel (Figure 4-2C). Figure 4-2D shows the locations of the two motifs examined by SPR within the PBM profiles, adding further confirmation of the impact of tandem T:A bps on binding. To summarize, the enhanced binding affinity of c-Rel homodimers in comparison to RelA homodimers allows it to bind to DNA motifs that show little resemblance to previously described NF-κB motifs.

Identification of Novel, Unrecognizable NF-κB in the *Il12b* Promoter

Unexpectedly, an examination of the mouse and human *Il12b* promoter sequences revealed that two conserved sequences with four tandem T:A bps (Figure 4-2E, NF-κB3 and NF-κB4) are located immediately downstream of two previously described non-consensus NF-κB motifs (Figure 4-2E, *Il12b* NF-κB1 and NF-κB2; Sanjabi et al., 2005). The *Il12b* NF-κB1 and NF-κB2

motifs, but not the *Il12b* NF-κB3 or NF-κB4 motifs, contain two or three tandem G:C bps, which are known to be critical for stable binding, in one half-site. Although the NF-κB3 and NF-κB4 motifs are unrecognizable as NF-κB motifs based on prior studies, our PBM and SPR results combined with their conservation through evolution (Figure 4-2E and below) led us to consider the possibility that they may contribute to c-Rel-dependent *Il12b* activation.

We first asked whether the NF-κB3 motif (shown below to bind c-Rel homodimers much more strongly than the NF-κB4 motif) contributes to *Il12b* transcriptional induction. This was first assessed using *Il12b* promoter-luciferase reporter plasmids containing the WT *Il12b* promoter and promoters with substitution mutations in the NF-κB1, NF-κB2, and NF-κB3 motifs. As previously shown (Murphy et al., 1995), c-Rel overexpression in HEK 293 cells activated the WT *Il12b* promoter much more strongly than RelA overexpression (Figure S4-2). Mutations in each of the three NF-κB motifs reduced transactivation by c-Rel in this assay (Figure S4-2).

To test the function of the NF-κB3 motif in a chromosomal context, we took advantage of a bacterial artificial chromosome (BAC) strategy developed in our lab prior to the emergence of CRISPR-Cas9 in mammalian cells (Figure S4-3; Doty et al., in preparation). Using a 200 kb BAC spanning the mouse *Il12b* locus, into which we had incorporated a green fluorescence protein cDNA (*Gfp*) to monitor BAC-derived *Il12b* expression, we introduced substitution mutations into the NF-κB1, NF-κB2, and NF-κB3 motifs (Figure S4-3B). Two additional mutations were examined as controls. The WT and mutant *Il12b-Gfp* BACs were then stably transfected into mouse embryonic stem cells (ESCs). After the selection of multiple clones for each BAC containing single-copy BAC integrants, the clones were differentiated in vitro into terminally differentiated macrophages, followed by LPS stimulation for 2 hr (Figure S4-3A). qRT-PCR was then used to monitor transcriptional induction of the *Il12b-Gfp* gene (using *Gfp* primers). Transcription of the endogenous *Il12b* gene (using *Il12b* primers), which should be activated similarly in all WT and mutant lines, was also examined by qRT-PCR. In this context, all three NF-κB motifs were found to be critical for transcriptional induction of the *Il12b-Gfp* BAC, whereas

the two control mutations (in a downstream NFAT motif and a non-conserved motif) had much smaller effects (Figure 4-2F).

c-Rel Homodimer Binding to the Tandem *Il12b* Promoter Motifs

Electrophoretic mobility shift assay (EMSA) titration experiments with a recombinant protein containing the c-Rel RHR (expressed in E. coli) revealed that the c-Rel homodimers can bind all four of the non-consensus motifs from the *Il12b* promoter, albeit with highly variable relative affinities (Figure 4-3A, B). SPR experiments revealed the different binding off-rates to the four motifs (Figure 4-3C), with NF-κB4 unable to bind with sufficient affinity for reliable results (data not shown). Recombinant RelA dimers bound poorly to all four sequences (Figure 4-3C). However, greatly increased off-rates were observed with the chimeric RelA(C46) protein that contains the 46 key residues of c-Rel in the context of the RelA protein (Figure 4-3C; see also Figure S4-4A for EMSA results with the chimeric protein).

A comparison of seven recombinant homodimeric and heterodimeric species - c-Rel:c-Rel, RelA:RelA, p50:p50, c-Rel:p50, RelA:p50 and RelB:p50 - revealed much stronger binding by c-Rel:c-Rel homodimers to both the NF-κB3 and NF-B4 motifs than by any of the other dimers (Figure S4-4B, C). In addition, an oligonucleotide containing substitution mutations in the tandem T:A bps of the NF-κB3 oligonucleotide exhibited greatly reduced binding in comparison to the WT NF-κB3 sequence (Figure 4-3D).

Importantly, in EMSA titration experiments using a radiolabeled probe containing the native NF-κB1, NF-κB2, and NF-κB3 sequences from the *Il12b* promoter, efficient and simultaneous binding of three c-Rel homodimers was observed (Figure 4-3E). Mutation of any of the three motifs eliminated the slowest mobility band (Figure S4-5A). This property was unique to c-Rel homodimers, as only one RelA homodimer was capable of binding the same oligonucleotide, presumably due to the much lower affinity of this homodimer for each of the three motifs (Figure 4-3E). p50:p50, p50:c-Rel, and p50:RelA dimers were also incapable of loading three dimers onto this oligonucleotide (Figure 4-3E, right). Importantly, four c-Rel homodimers could simultaneously

bind a radiolabeled probe containing all four potential recognition sequences from the *Il12b* promoter (Figure S4-5B).

A careful examination of the relative abundances of complexes containing one, two, three, or four bound dimers suggested the possibility of cooperative binding. To examine this possibility, we performed gel shift titrations with different combinations of two of the tandem motifs. Cooperative binding was not observed between the NF-κB1 and NF-κB2 motifs or between the NF-κB3 and NF-κB4 motifs (Figure S4-6). However, strong cooperativity between the NF-κB2 and NF-κB3 motifs was observed, in that the efficiency of c-Rel homodimer binding to NF-κB3 was greatly enhanced when the probe contained an adjacent NF-κB2 motif (Figure S4-6).

Together, these results add strong support for the hypothesis that preferential binding of c-Rel homodimers to three and possibly four tandem recognition motifs in the *Il12b* promoter underlies the c-Rel dependence of *Il12b* transcription. Moreover, the *Il12b* NF-κB3 and NF-κB4 motifs would have been entirely unrecognizable on the basis of past knowledge of NF-κB dimer binding specificities.

c-Rel Versus RelA Preferential Binding In Vivo

To determine whether preferential binding of c-Rel can be observed in vivo, we performed ChIP-seq in BMDMs stimulated with lipid A for 0 and 60 min. From an analysis of three biological replicates with c-Rel, RelA, and p50 antibodies, 7,257, 4,595, and 1,906 reproducible called peaks, respectively, were observed. The widely different numbers of peaks are likely to be a reflection primarily of variable antibody qualities but may also be influenced by differences in genomic interactions by dimers containing each of the three proteins.

To examine the possibility of selective binding by c-Rel or RelA, we selected the strongest 6,700 peaks obtained with each antibody, which upon integration of data sets yielded 8,134 distinct ChIP-seq peaks (Figure 4-4A). We then determined their RPKM ratio at each of these 8,134 genomic locations. This analysis revealed that the median c-Rel RPKM was 1.89-fold greater than the RelA RPKM, most likely due to differences in antibody quality (Figure 4-4A, right).

7,549 peaks exhibited a c-Rel/RelA RPKM ratio between 1.3 and 3, with 113 peaks exhibiting a ratio >3 and 472 peaks exhibiting a ratio <1.3 (Figure 4-4A, right).

To determine whether selective binding by either c-Rel or RelA is influenced by binding motifs, we performed de novo DNA binding motif analysis with the 300 peaks exhibiting the largest and smallest RPKM ratios (top and bottom 4% of all peaks). Strikingly, this analysis revealed that the peaks exhibiting the largest c-Rel/RelA RPKM ratios showed the strongest enrichment for motifs resembling an NF-κB half-site (Figure 4-4B, GGAA) with two G:C bps followed by two A:T bps. The enrichment of a half-site is consistent with the hypothesis that the enhanced binding affinity of c-Rel homodimers observed in vitro allows preferential c-Rel binding to relatively low-affinity sites in vivo. It may be noteworthy that c-Rel preferential binding may occur at a diverse range of non-consensus sites that do not reveal a defined enriched motif in a conventional de novo motif analysis. Notably, a consensus NF-κB motif was not among any of the four most enriched motifs exhibiting preferential c-Rel binding in this analysis.

In contrast to the observations with c-Rel-preferential peaks, RelA-preferential peaks exhibited the strongest enrichment of motifs resembling RelA:p50 or c-Rel:p50 heterodimers, with three tandem G:C bps at one half-site (optimal for p50 interactions) and two tandem G:C bps at the other half-site (optimal for RelA or c-Rel interactions), with five intervening bps (Figure 4-4B). Enrichment of this conventional heterodimer motif suggests that the RelA interactions at sites exhibiting the lowest c-Rel/RelA binding ratios are mediated by RelA:p50 heterodimers. This finding could be a reflection of a high abundance of RelA:p50 heterodimers in comparison to c-Rel:p50 heterodimers or noise in the analysis. Alternatively, RelA:p50 interactions with a small subset of consensus motifs may be selectively stabilized by other proteins that interact specifically with RelA (Wan et al., 2007).

For further comparison, we performed a de novo motif analysis with 300 peaks from the middle of the c-Rel/RelA RPKM ratio spectrum. Here, we observed strong enrichment of a near-consensus motif that contains two clear NF-κB half-sites. The composition of this enriched motif suggests that it could represent a composite of interactions by a variety of dimeric species,

including RelA:p50, c-Rel:p50, and c-Rel:RelA heterodimers, and possibly c-Rel:c-Rel and RelA:RelA homodimers.

We next examined how the enrichments of four distinct NF-κB consensus motifs (three motifs from the Homer program and one de novo motif from the above analysis) change across the c-Rel/RelA RPKM ratio spectrum, after dividing all 8,134 peaks into 27 bins. This analysis revealed a consistent trend toward weaker enrichment of all four consensus motifs as the c-Rel/RelA RPKM ratio increases (Figure 4-4C). This is consistent with the view that c-Rel homodimers bind a broader range of non-consensus sequences.

Most importantly, we graphed c-Rel/RelA RPKM ratios for 1,120 peaks that annotate to promoter regions (-1kb to +1kb relative to the transcription start site [TSS]) (Figure 4-4D). This analysis revealed strongly preferential c-Rel binding to the promoters of three genes that exhibited strong c-Rel-dependent transcription: *Il12b*, *Tnfsf9*, and *Noct*, with more limited preferential binding at the *Il4i1* promoter and at two other peaks that annotate to the *Tnfsf9* and *Noct* promoters (Figure 4-4D). Thus, these ChIP-seq results add support to the structural and biochemical evidence of NF-κB dimer-specific binding preferences, with evidence that these binding preferences are relevant to the regulation of the *Il12b* gene and at least a few other c-Rel-dependent genes.

c-Rel Versus p50 Preferential Binding In Vivo

We next used the same approach as above to examine preferential binding by c-Rel in comparison to p50. Because of the smaller number of p50 peaks, we limited this analysis to the top 2,891 peaks obtained with each antibody, which, when merged, allowed us to include 4,414 peaks in the analysis (Figure 4-5A). An analysis of c-Rel/p50 RPKM ratios showed that c-Rel peaks were generally stronger, most likely again reflecting higher quality c-Rel antibodies (Figure 4-5A).

De novo motif analysis performed with 300 peaks showing the largest c-Rel preference (top 6.7% of all peaks) revealed strong enrichment of two types of motifs (Figure 4-5B): The most

enriched motifs were characteristic of c-Rel or RelA homodimer binding, with two G:C bps within each half-site. Strong enrichment was also observed with motifs characteristic of an NF-κB half-site (Figure 4-5B). Interestingly, motifs showing the smallest c-Rel/p50 RPKM ratios exhibited strong characteristics of p50 homodimer binding. The most enriched motifs contained three G:C bps within each half-site (separated by five bp), consistent with p50 interactions with three G:C bps (Figure 4-5B). Strong enrichment was also observed with a motif containing five tandem G:C bps, which perfectly matches a de novo motif that exhibited the greatest enrichment in prior PBM experiments among oligonucleotide sequences with p50 homodimer preferential binding in comparison to RelA and c-Rel homodimers (Siggers et al., 2011).

An analysis of enriched motifs across the full spectrum of c-Rel/p50 RPKM ratios revealed that all three NF-κB motifs from the Homer program exhibit their lowest enrichment among c-Rel preferential peaks, most likely reflecting once again the ability of c-Rel homodimers to bind divergent motifs and half-sites (Figure 4-5C). The greatest differences between p50-preferential and c-Rel-preferential peaks were observed with two of the de novo motifs defined above (Figure 4-5B). (Motif enrichment with NF-κB half-site motifs was uninformative because half-sites are also contained within dimeric motifs.)

Notably, a p50 ChIP-seq peak was not observed at the *Il12b* promoter (data not shown). The absence of a p50 peak is consistent with our evidence that the *Il12b* promoter is regulated by a c-Rel homodimer. However, because of the small overall number of p50 ChIP-seq peaks, we cannot exclude the possibility that the absence of a p50 peak is due to relatively poor p50 antibody quality.

To summarize, these ChIP-seq studies provide strong confirmation that the DNA binding characteristics of various NF-κB dimers observed in biochemical experiments accurately reflect dimer-specific differences in vivo at a genome-wide scale. In particular, c-Rel preferential binding occurs at genomic locations that often lack known NF-κB consensus sequences recognized by heterodimeric species (RelA:p50 and c-Rel:50) or by p50 homodimers. Of greater importance, c-Rel preferential binding in vivo was observed at the *Il12b* promoter and at the promoters of a

subset of other genes that exhibit relatively strong c-Rel-dependence in macrophages, providing a strong connection between c-Rel-dependent transcriptional induction and the intrinsic binding differences between c-Rel and RelA dimers.

Late Emergence of the c-Rel-RelA DNA Binding Difference During Vertebrate Evolution

The above results suggest that the duplication and divergence of the genes encoding c-Rel and RelA coincided with the emergence of adaptive immunity and that, even in an innate immune cell type, a critical role of c-Rel is to regulate a gene, *Il12b*, whose primary function is to regulate adaptive immune responses. Furthermore, our findings strongly suggest that c-Rel-dependent induction of *Il12b* transcription is due at least in part to divergent DNA-binding affinities of c-Rel and RelA. Together, these results suggest that the evolutionary divergence of c-Rel and RelA in the earliest vertebrates may have been driven by mutations that resulted in the DNA-binding affinity differences, thereby facilitating neofunctionalization of c-Rel dimers.

To test this hypothesis, we isolated the RHRs of c-Rel and RelA from elephant shark (cartilaginous fish), zebra fish (bony fish), frogs (amphibians), chicken, and humans. The proteins were expressed in HEK 293 cells and extracts were prepared. Competition time course experiments were then performed by EMSA to measure relative binding affinities of the two proteins in each species (Figure 4-6A). For these experiments, the proteins were pre-bound to a radiolabeled oligonucleotide containing a consensus NF-κB motif. Then, a large excess of an unlabeled oligonucleotide containing the same sequence was added and the samples were loaded onto a polyacrylamide gel at different times. When the bound protein dissociates from the radiolabeled oligonucleotide, re-binding will almost always be to the excess unlabeled oligonucleotide, such that the decline is radiolabeled protein-DNA complex provides a measure of protein off-rate and relative binding affinity.

Using this approach, the large off-rate (affinity) difference between mouse c-Rel and RelA was readily observed, as RelA binding to the labeled oligonucleotide declined substantially one minute after the addition of the unlabeled oligonucleotide (Figure 4-6A, B). In contrast, the c-Rel

complex was stable throughout the 20-minute time course. A similar off-rate difference was observed with the human and frog orthologs of c-Rel and RelA. Surprisingly, however, RelA and c-Rel proteins from elephant shark, zebrafish, and chicken exhibited similar off-rates (Figure 4-6A, B). These results suggest that, in contrast to the above hypothesis, the initial evolutionary divergence of c-Rel and RelA was not driven by the key DNA-binding affinity difference. Instead, the divergence was likely supported by another difference that remains to be defined (see below).

Additional insights emerge from a comparison of the *Il12b* promoter sequences in various vertebrate species (Figure 4-6C). The four potential *Il12b* NF-κB motifs are well-conserved among placental mammals. The *Il12b* NF-κB1, NF-κB2, and NF-κB3 are well-conserved in marsupials (platypus and opossum), with NF-κB4 also significantly conserved in platypus. However, in the chicken genome, only the *Il12b* NF-κB1 and NF-κB2 motifs display conservation, but only to a limited extent. The *Il12b* promoter sequences from earlier vertebrates do not align with the mammalian sequences. These results add support to the notion that the divergent DNA-binding properties of c-Rel and RelA that are responsible for the strong c-Rel-dependence of *Il12b* transcription in mammals emerged at a relatively late stage of vertebrate evolution.

DISCUSSION

We combined structural, biochemical, genomic, functional, and evolutionary analyses to examine a key difference between the c-Rel and RelA NF-κB paralogs. PBM and SPR experiments revealed previously unknown differences in binding preferences, and X-ray crystallographic analysis provided insight into the structural underpinnings of these differences, which exist despite identical DNA-contacting residues in the two proteins. Biochemical and functional analyses showed that c-Rel homodimers can selectively bind four tandem motifs in the promoter of the *Il12b* gene, an important regulator of T-cell differentiation that exhibits strong c-Rel-dependence in activated macrophages. Two of these NF-κB motifs are similar to previously unrecognizable NF-κB binding motifs identified by PMB, with cooperative binding of c-Rel homodimers to two of the *Il12b* promoter motifs possibly contributing to the selectivity mechanism. ChIP-seq

experiments demonstrated that the distinct binding properties of c-Rel, RelA, and p50 dimers defined biochemically can largely explain their preferential distributions across the genome. Finally, despite the key c-Rel and RelA homodimer binding differences and the strong evidence that these differences are functionally important, they were not observed with RelA and c-Rel orthologs from early vertebrates, suggesting that the initial divergences of the two genes were supported by other differences.

The absence of intrinsic DNA-binding differences between RelA and c-Rel in early vertebrates raises the question of what other properties of the two proteins and their genes drove their divergence and specialization. One likely contributor to their specialization is their differential expression. The differential developmental expression of the *Rela* and *Rel* genes is well-known (Gilmore and Gerondakis, 2011), with *Rela* expressed at a high level much more broadly than *Rel*. Differences in stimulus-responsive nuclear translocation of RelA and c-Rel complexes have also been described (DeLaney et al., 2019). These expression differences may have allowed the expansion of NF-κB's regulatory potential.

Differences in the RelA and c-Rel C-terminal transactivation domains may have also contributed to the divergence and specialization of the two family members, possibly contributing to their regulation of distinct sets of genes in different settings. Although our prior chimeric protein studies showed that the C-terminal transactivation domains or RelA and c-Rel and interchangeable for the activation of the endogenous *Il12b* gene in *Rel*$^{-/-}$ macrophages, RelA's transactivation domain carries out a well-documented interaction with the p300/CBP co-activators (Zhong et al., 1998; Hayden and Ghosh, 2012). This interaction does not appear to be supported by c-Rel's transactivation domain (Wang et al., 2007). The p300/CBP interaction of RelA has been shown to be critical for the activation of a distinct subset of NF-κB target genes (Dong et al., 2008), suggesting that c-Rel may not be capable of activating these genes and suggesting that this RelA/c-Rel difference may have contributed to the initial divergence of the two genes.

The late evolution of the DNA binding difference between RelA and c-Rel suggests that this intrinsic biochemical difference may have been necessary to support immunoregulatory

functions not required in the earliest vertebrates. This may include contributions to fetal tolerance in placental mammals and/or tolerance mechanisms associated with metamorphosis in amphibians. The finding that the DNA-binding differences between c-Rel and RelA are apparent in the frog proteins, but not the chicken proteins, is interesting to consider. Since amphibians and birds share an evolutionary ancestor, the DNA-binding differences may have emerged at an earlier stage of vertebrate evolution and may have then been lost in some or all bird species.

Consistent with the possibility that divergent DNA-binding properties of RelA and c-Rel may have been important for the evolution of advanced tolerance mechanisms, c-Rel is known to play critical selective roles in Treg development and function (Isomura et al., 2009). An enhancer for the *Foxp3* gene has been identified that evolved late during vertebrate evolution, with a hypothesized role in the support of fetal tolerance (Samstein et al., 2012). Amino acid changes in the FoxP3 protein that are specific to mammals have also been described (Andersen et al., 2012).

Whether the DNA-binding differences described here are critical for the distinct activities of RelA and c-Rel in Treg development and function remains to be determined (Zheng et al., 2010; Oh et al., 2017). It also is not yet known whether the binding differences are important for the distinct functions that have been documented for RelA and c-RelA in other cells of the adaptive immune system. In support of this possibility that the DNA-binding differences are important in other cell types, the first DNA motif proposed to confer c-Rel-selectivity in T cells is located in a control region for the *Il2* gene, referred to as the CD28 response element (Shapiro et al., 1997). Although both RelA- and c-Rel-containing dimers can bind this motif (Kane et al., 2002; Tuosto, 2011). it possesses a non-consensus sequence that would be predicted to allow preferential activation by c-Rel homodimers.

Another unanswered question is why *Il12b* acquired a critical requirement for c-Rel for its activation in macrophages, whereas many other putative NF-κB target genes appear to be activated redundantly by RelA and c-Rel complexes. This finding suggests that *Il12b* would not function properly if it were regulated similarly to other NF-κB target genes, thereby mandating its requirement for c-Rel. A likely reason for the emergence of a c-Rel requirement would be that c-

Rel dimers are subject to unique regulatory mechanisms. Although dimers containing c-Rel and RelA are thought to be maintained in an inactive state by the same IκBα and IκBβ proteins, activation differences have occasionally been reported, including a unique requirement for Caspase 8 for c-Rel nuclear translocation in LPS-stimulated macrophages (DeLaney et al., 2019). Proper regulation of *Il12b* expression in physiological settings may therefore rely on this and other unique mechanisms for regulating the nuclear translocation of c-Rel dimers.

RelA and c-Rel expression kinetics are also regulated differently in activated macrophages, with c-Rel expression more prolonged than RelA expression in response to some stimuli (York et al., personal communication). This difference can allow c-Rel dimers to play a broader role in inflammatory gene regulation after RelA dimers are downregulated. These RelA/c-Rel expression differences would allow a broad range of inflammatory genes to be sensitive to the modulation of c-Rel activity, as recently described for the regulation of c-Rel activity by very long chain ceramides associated with anti-inflammatory pathways (York et al., submitted).

Finally, the absence of the intrinsic DNA-binding difference between RelA and c-Rel homodimers in early vertebrates raises the question of whether *Il12b* induction in early vertebrates remains c-Rel-dependent, but with the c-Rel dependence due to a different mechanism, or whether *Il12b* induction in these species is perhaps regulated by other mechanisms. Given the poor conservation of the *Il12b* promoter in these species, the involvement of a different regulatory mechanism may not be surprising. Possibly related to this question, *Il12b* expression in mammals appears to be c-Rel independent in some settings (Grumont et al. 2001), suggesting the existence of a pathway capable of circumventing the c-Rel requirement.

EXPERIMENTAL PROCEDURES

Cloning of NF-κB cDNAs from Sea Lamprey

mRNA was prepared from the whole blood of lamprey obtained from the Great Lakes (species?). After preparing cDNA, gene segments encoding NF-κB family member RHRs were amplified by PCR using primers designed on the basis of the partial lamprey genome sequence, as well as

degenerate synthetic primers designed on the basis of NF-κB RHR sequences from various vertebrate species. This approach led to the successful isolation of five distinct RHR-encoding genes, including genes not yet included in reported lamprey genomes, which remain incomplete.

Expression of NF-κB RHRs from Vertebrate Species

Vertebrate RHR cDNAs were amplified by PCR and cloned into a pcDNA3 vector with an N-terminal Flag epitope tag. c-Rel and RelA RHR cDNAs were from mouse (Sanjabi et al., 20025), human, chicken, frog (*Xenopus laevis*), zebrafish, and elephant shark. The recombinant proteins were expressed in HEK 293T cells following transient transfection of the expression plasmids. Nuclear extracts were prepared for EMSA as previously described (Plevy et al., 1997; Sanjabi et al., 2005).

X-Ray Crystallography

The RelA(C46) RHR protein was expressed and purified as described (Siggers et al., 2011). Crystals were then prepared (after numerous tests of conditions) after pre-binding to the same double-stranded oligonucleotide sequence (with an NF-κB consensus motif) used to solve the co-crystal structure of mouse RelA homodimers bound to DNA. X

PBM and SPR Experiments

The PBM experiments were described previously (Siggers et al., 2011), with further computational scrutiny of bound oligonucleotides revealing frequent binding to oligonucleotides containing the tandem T:A bps. SPR was performed as described (Siggers et al., 2011), using purified recombinant RHRs as described (Siggers et al., 2011).

Transient Transfection Experiments

The mouse *Il12b* promoter (-355 to +55) was cloned into the pGL4.10 vector (Promega). Motif mutations were introduced using the GENEART site-directed mutagenesis system (Invitrogen).

C-terminal single Flag tag versions of full-length cDNAs encoding mouse RelA, c-Rel, and p50 were cloned into pcDNA3 vectors. 293T cells, grown in DMEM with 10%FBS, were transfected with Lipofectamine 2000 (Invitrogen) in 24-well plates. 1-8,000 ng of the expression plasmid were co-transfected with 20ng of the *Il12b*-pGL4 vectors. 1 ng of a TK-Renilla luciferase vector was included as a transfection control. Cells were collected 24-hr post-transfection and firefly and renilla luciferase were analyzed with the Dual Luciferase Reporter Assay system (Promega). Normalization was performed using Western blot analysis for protein concentration, renilla luciferase for transfection efficiency, and empty vector firefly/renilla values for background signal.

Engineering and Analysis of Mutant *Il12b-Gfp* BACs

Recombination in E. coli by standard procedures (Warming et al., 2005) was used to introduce a *Gfp* cDNA into the second exon of a 200 kb BAC spanning the mouse *Il12b* locus (Doty et al., unpublished results). Recombination in E. coli was then used to introduce substitution mutations into *Il12b* promoter motifs (see Figure S4-3B). The recombinant WT and mutant *Il12b-Gfp* BACs were introduced into a feeder-free mouse ESC line. Clonal lines were expanded and single-copy integrants were identified by qPCR and Southern blot analysis (Figure S4-3 and data not shown).

BAC-containing ESCs were differentiated into macrophages as described (Keller, 2002), with confirmation of differentiation into myeloid progenitors and macrophages confirmed by flow cytometry (Figure S4-3A). Macrophages were then stimulated with lipid A (100ng/ml) for 2 hr, followed by qRT-PCR analysis of BAC-derived *Gfp* mRNA. Endogenous *Il12b* mRNA was monitored in each line by qRT-PCR as a positive control.

Electrophoretic Mobility Shift Assays (EMSAs) and Immunoblot Assays

Purified proteins (Figures 4-3 and S4-1-S4-4) or nuclear extracts from transfected HEK293T cells (Figure 4-6) were incubated with radiolabeled probe (in 25μL total volume) in the presence of 10mM Tris-HCl, 150mM NaCl, 1mM DTT, 1mM EDTA, 5% glycerol, 80ng/μL dI-dC, and 200ng/μL BSA, with limited amounts of double-stranded ^{32}P-labeled probes (~10^{-11} M). Each reaction was

incubated at 4°C for 30 minutes before gel electrophoresis. Gel shift assays were run as described (Plevy *et al.*, 1997). Band intensities were quantified using ImageQuant (GE).

Immunoblots were performed as previously described (Siggers et al., 2011), using antibodies directed against RelA or c-Rel.

Nascent Transcript RNA-seq Experiments

Bone marrow-derived macrophages were prepared from C57BL/6 and *Rel*$^{-/-}$ mice as described (Tong et al., 2016). Macrophages were activated on day 6 with lipid A for the time indicated in each experiment (Sigma-Aldrich). Nascent transcript RNA-seq was performed and the data were analyzed as described (Bhatt et al., 2012; Tong et al., 2016).

ChIP-Seq Analysis

ChIP-seq was performed as previously described (Barish et al., 2010; Tong et al., 2016) with anti-RelA, anti-c-Rel, and anti-p50 antibodies (Cell Signaling, Inc. 8242, 68489, and 13586, respectively). Approximately 10 million BMDMs were used per sample from C57BL/6 mice aged 8-12 weeks. After crosslinking with 1mM DSG and 1% PFA, cells were sonicated on a Covaris M220 focused ultrasonicator to 200-500-bp DNA fragments. ChIP-seq libraries were prepared and peaks were called as described (Heinz et al., 2010; Tong et al., 2016). Read were aligned using Hisat2 to the NCIB37/mm9 mouse genome. To compare peaks across multiple samples, a master probe was generated with BED Tools (Quinlan et al., 2010).

FIGURE LEGENDS

Figure 4-1. NF-κB Phylogenetic Analysis and Highly Selective c-Rel Requirement

(A) A phylogenetic tree was prepared with the RHR amino acid sequences of NF-κB family members from representative invertebrate and vertebrate species (see Methods). All sequences were from public databases, except the sea lamprey sequences were obtained from cDNAs

isolated by degenerate PCR from mRNA in freshly isolated sea lamprey blood. Only a subset of the sea lamprey NF-κB genes are in reported genomes.

(B) Volcano plots were prepared for a comparison of nascent transcript RNA-seq datasets from lipid A-induced genes in WT versus *Rel*$^{-/-}$ C56BL/6 macrophages. Comparisons of WT versus mutant RPKM are from data at 60- and 120-min post-stimulation. The genes exhibiting the strongest c-Rel dependence and weaker, statistically significant dependence are highlighted in red and gold, respectively. The identities of genes exhibiting the strongest c-Rel-dependence at each time point are included, along with their fold-induction in WT macrophages at that time point (in parenthesis).

Figure 4-2. PBM and SPR Evidence of NF-κB Binding to Novel Motif Sequences

(A) A consensus recognition motif representative of the preferred binding sites for both RelA and c-Rel homodimers determined by PBM is shown (Siggers et al., 2012).

(B) PBM scatter plots comparing the binding (in arbitration units) of recombinant RelA (x-axis) and c-Rel (y-axis) homodimers to a broad range of double-stranded oligonucleotide sequences are shown. Oligonucleotides containing the sequence shown in each graph are highlighted in red. These graphs reveal binding to many oligonucleotides containing the sequence, TTTT. By comparison, the proteins generally bind more weakly to oligonucleotides containing the sequences, GAGAT and CACTA, and to oligonucleotides containing, the sequence, GGGGG, despite the fact that RelA and c-Rel half-sites typically include two tandem G:C bps. The diagonal shape shows the RelA and c-Rel homodimers have similar spectrums of DNA preferences, but the relative affinities of the two dimers for each motif cannot be determined from PBM data.

(C) SPR analysis reveals that c-Rel homodimers bind a double-stranded oligonucleotide sequence, GGGGGTTTTT, with a much slower off-rate than the sequence, GGGGGGAGAT, further supporting a specific influence of the sequence, TTTTT, within one half-site. Off-rates are displayed as the time (sec) needed for half of the protein to dissociate from the probe. RelA homodimers bind the sequence, GGGGGTTTTT, with a much faster off-rate than c-Rel

homodimers, consistent with prior evidence that the intrinsic DNA-binding affinity of c-Rel homodimers (to all motifs) is much greater than that of RelA homodimers. The RelA (C46) protein binds with a much slower off-rate than RelA homodimers, confirming the importance of the 46 c-Rel residues for higher-affinity binding.

(D) The locations of the oligonucleotide sequences shown in panel C are shown on PBM profiles comparing either RelA versus c-Rel homodimers or comparing c-Rel versus RelA (C46) homodimers.

(E) The sequences of portions of the mouse and human *Il12b* promoters are shown, with conserved nucleotides indicated (asterisks). The previously described non-consensus c-Rel homodimer binding sites (NF-κB1 and NF-κB2) and a previously described binding site for C/EBP proteins are indicated. Two additional conserved, potential NF-κB binding motifs (NF-κB3 and NF-κB4) are indicated. The GTTTT sequences that may help support NF-κB binding are highlighted in red.

(F) A BAC encompassing the mouse *Il12b* locus was engineered in E. coli to contain Gfp-expressing sequences, and was then further engineered with substitution mutations in the *Il12b* NF-κB1, NF-κB2, or NF-κB3 sequences (Figure S2), with mutations introduced into a downstream NFAT sequence and a non-conserved sequence in the *Il12b* promoter as controls (see Figure 6C). The recombinant BACs were then stably integrated into mouse ESC and multiple clones containing single-copy BAC integrants were selected for each mutant. Following differentiation of the ESC into terminally differentiated macrophages, cells were stimulated with LPS for 0 or 2 hrs, followed by mRNA isolation. qRT-PCR was then used to quantify relative mRNA levels in each clone for the endogenous *Il12b* mRNA and the BAC *Il12b-Gfp* mRNA (detected using *Gfp* primers). The graph displays mean relative mRNA levels and standard errors as a percentage of *Gapdh* mRNA level (y-axis, log10 scale) from 2-6 independent clones for the WT BAC clone and each BAC mutant.

Figure 4-3. c-Rel Homodimer-Specific Binding to the *Il12b* Promoter Motifs

(A) EMSA experiments were performed with radiolabeled double-stranded oligonucleotides containing each of the four potential non-consensus NF-κB motifs from the mouse *Il12b* promoter. Increasing concentrations of recombinant c-Rel RHR homodimer protein expressed in E. coli were used. The locations of the predicted homodimer-DNA complexes are indicated (arrowheads), along with complexes that may represent aggregates formed at high protein concentrations (asterisks).

(B) The line graph shows the percentage of each radiolabeled oligonucleotide probe from panel A bound by c-Rel protein at each protein concentration. The table (right) shows the concentration of c-Rel protein (nM) required to bind 50% of the radiolabeled probe.

(C) SPR was used to determine off-rates of c-Rel, RelA, and RelA (C46) homodimers from the *Il12b* NF-κB1, NF-κB2, and NF-κB3 motifs. No reproducible binding was observed with the NF-κB4 sequence.

(D) EMSAs were used to examine the binding of increasing concentrations of the recombinant c-Rel RHR protein to radiolabeled, double-stranded oligonucleotide probes containing the WT NF-κB3 motif and a motif with multiple nucleotide substitutions (κB3M).

(E) EMSAs were used to examine binding of purified recombinant c-Rel:c-Rel, RelA:RelA, p50:p50: c-Rel:p50, and RelA:p50 dimers to a radiolabeled oligonucleotide probe containing the Il12b promoter sequence spanning the NF-κB1, NF-κB2, and NF-κB3 motifs. The locations of complexes containing one, two, or three bound dimers are indicated. The unbound radiolabeled probe is visible at the bottom of the image.

Figure 4-4. Selective Binding of c-Rel and RelA in Mouse BMDMs Examined by ChIP-seq

(A) ChIP-seq experiments were performed with antibodies to RelA and c-Rel in BMDMs stimulated with lipid A for 0 and 1 hr. The xxx peaks with the strongest peak scores obtained with each antibody at the 1-hr time point were selected and merged, yielding 8,134 peaks that yielded a strong peak score with one or both proteins. c-Rel/RelA ratios were then calculated from the RPKMs at each peak and were plotted (ratios on the y-axis), with the 8,134 peaks along the x-

axis (displayed as a percentage of total peaks). The top and bottom 2% of peaks based on the ratio (163 peaks) were then selected for motif analysis. At the right, numbers of peaks with different ranges of RPKM ratios are shown.

(B) The most enriched motifs from a de novo motif analysis performed with Homer are shown. Motif analysis was performed with 163 peaks (2%) representing the largest peak ratios, 163 peaks representing the smallest peak ratios, and 163 peaks from the middle of the ratio distribution. The peaks with the largest c-Rel/RelA RPKM ratios exhibit the strongest enrichment of motifs resembling NF-κB half-sites and may have missed other more divergent non-consensus motifs capable of binding c-Rel. The peaks with the smallest c-Rel/RelA RPKM ratios exhibit the strongest enrichment of motifs resembling a RelA:p50 or c-Rel:p50 heterodimer consensus, with three G:C bps in one half-site (preferred by p50) and two G:C bps in the other half-site (preferred by RelA or c-Rel). The peaks in the middle exhibit enrichment of motifs that resemble the NF-κB consensus and may represent a hybrid between c-Rel:50 (or RelA:p50) motifs and c-Rel:c-Rel (or RelA:RelA) motifs.

(C) The 8,134 peaks were divided into 27 bins of equal size across the spectrum of c-Rel/RelA peak ratios. The enrichment of three consensus NF-κB motifs within the Homer program, and the consensus NF-κB motif defined in panel B, was then examined across the spectrum. All four motifs exhibited greater enrichment in bins exhibiting the lowest c-Rel/RelA RPKM ratios, consistent with the notion that c-Rel-preferential binding occurs primarily at non-consensus sequences.

(D) Among the 8,134 peaks, 1,120 annotate to promoter regions (-1kb-+1kb relative to the TSS of annotated genes). The c-Rel/RelA RPKM ratios of these peaks were plotted and used to examine the relative ratios of peaks within the promoters of genes that exhibited strong c-Rel-dependence in Figure 1B. The *Il12b* promoter peak and one of two c-Rel/RelA peaks within the *Tnfsf9* promoter ranked among the top 1.25% of peaks with the largest c-Rel/RelA RPKM ratios. Lesser c-Rel preferences were observed at a second peak in the *Tnfsf9* promoter, at two peaks that annotated to the *Noct* promoter, and at a peak in the *Il4i1* promoter. The ratio rank of each of

these six peaks, the percentile among all promoter peaks, and the c-Rel/RelA RPKM ratio (FC) are shown at the right.

Figure 4-5. Selective Binding of c-Rel and p50 in Mouse BMDMs Examined by ChIP-seq

(A) ChIP-seq experiments were performed with antibodies to c-Rel and p50 in BMDMs stimulated with lipid A for 0 and 1 hr. The xxx peaks with the strongest peak scores obtained with each antibody at the 1-hr time point were selected and merged, yielding 4,414 peaks that yielded a strong peak score with one or both proteins (the peak number of smaller than in the RelA/c-Rel comparison because of the small number of peaks obtained with p50 antibody, probably due to relatively poor antibody quality). c-Rel/p50 ratios were then calculated from the RPKMs at each peak and were plotted (ratios on the y-axis), with the 4,414 peaks along the x-axis (displayed as a percentage of total peaks). The top and bottom 2% of peaks based on ratio (89 peaks) were then selected for motif analysis. At the right, numbers of peaks with different ranges of RPKM ratios are shown.

(B) The most enriched motifs from a de novo motif analysis performed with Homer are shown. Motif analysis was performed with 89 peaks (2%) representing the largest peak ratios, 89 peaks representing the smallest peak ratios, and 89 peaks from the middle of the ratio distribution. The peaks with the largest c-Rel/p50 RPKM ratios exhibit the strongest enrichment of two motifs, one resembling NF-κB half-sites and a second resembling sequences preferred in vitro for c-Rel or Rel A homodimers (or c-Rel:RelA heterodimers, with two G:C-bps in each half-site). The peaks with the smallest c-Rel/p50 RPKM ratios exhibit the strongest enrichment of motifs resembling a p50 homodimer preferential motif defined in vitro (with three G:C-bps in each half-site), or a motif with five tandem G:C-bps, which is remarkably similar to a p50-homodimer preferential peak identified by PBM (Siggers et al., 2012). The peaks in the middle exhibit enrichment of motifs that resemble a conventional NF-κB consensus or an NF-κB half-site.

(C) The 4,414 peaks were divided into 15 bins of equal size across the spectrum of c-Rel/p50 peak ratios. The enrichment of the three consensus NF-κB motifs within the Homer program, as

well as one of the p50-preferential motifs and one of the c-Rel-preferential motifs identified in panel B, was then examined across the spectrum. The four motifs that resemble either p50 homodimer binding motifs (three G:C-bps in each half-site) or a c-Rel:p50 (or RelA:p50) heterodimer binding motif (three G:C-bps in one half-site and two G:C-bps in the other half-site), exhibited decreasing enrichment as the c-Rel/p50 RPKM ratio increased, consistent with the view that c-Rel dimers, especially c-Rel homodimers, often bind non-consensus sequences. In contrast, the motif resembling a c-Rel (or RelA) homodimer binding motif (two G:C-bps in each half site) exhibited increasing enrichment as the c-Rel/p50 RPKM ratio increased. NF-κB half-sites were equally enriched across the spectrum, consistent with the presence of half-sites in all of the full consensus motifs examined.

Figure 4-6. Late Evolution of the c-Rel and RelA DNA-Binding Differences

(A) cDNAs encoding the RHRs of RelA and c-Rel from six vertebrate species (see Methods) were over-expressed in HEK293T cells. Nuclear extracts were prepared and EMSA off-rate experiments were performed with a radiolabeled, double-stranded oligonucleotide probe containing a consensus NF-κB motif. For these experiments, after pre-binding of the protein to the radiolabeled probe, a large excess of the same oligonucleotide without the radiolabel was added. At 0, 1, 5, 10, and 20 min after the addition of the unlabeled oligonucleotide, a sample was added to the native polyacrylamide gel. Since protein that releases from the radiolabeled probe is far more likely to re-bind the unlabeled oligonucleotide in excess, measurement of the rate of loss of the protein-DNA complex provides an approximate measure of its half-life.

(B) The half-lives of the protein-DNA complexes examined in panel A are shown. The results are representative of at least three independent replicates for each pair of paralogs.

(C) The sequence of the *Il12b* promoter is compared in multiple vertebrate species.

SUPPLEMENTARY FIGURE LEGENDS

Figure S4-1. Lamprey NF-κB RHR Sequences

RHR sequences for the five NF-κB family members isolated from sea lamprey blood mRNA are shown. Residues exhibiting conservation with NF-κB family members from other vertebrate species are highlighted.

Figure S4-2. Functional Analysis of Non-Consensus NF-κB Motifs in the Mouse *Il12b* Promoter by c-Rel and RelA Overexpression in HEK 293T Cells.

HEK 293T cells were transiently transfected with *Il12b* promoter-luciferase reporter plasmid and expression plasmids for full-length Flag epitope-tagged RelA, c-Rel, and p50. Luciferase activity was monitored with plasmids containing the WT *Il12b* promoter and plasmids containing substitution mutations in the *Il12b* NF-κB1, NF-κB2, and NF-κB3 motifs. As previously observed (Murphy paper), c-Rel overexpression activated *Il12b* promoter activity more strongly than RelA overexpression. Promoter activity was compromised to variable extents by disruption of each of the three motifs.

Figure S4-3. Functional Analysis of Non-Consensus NF-κB Motifs in the Mouse *Il12b* Promoter Using Recombinant BACs in ESC-Derived Macrophages.

(A) The strategy used to analyze substitution mutations in the mouse *Il12b* promoter (analysis performed prior to the emergence of CRISPR-Cas9 in mammalian cells) is shown. Recombinant WT and mutant BACs spanning the *Il12b* locus (Doty et al., in preparation) were integrated into mouse ESCs, which were then differentiated into terminally differentiated macrophages that support pro-inflammatory responses following LPS stimulation comparable to BMDMs (not shown). The ESC-derived macrophages were then stimulated for 2 hr with LPS. Flow cytometry was used to monitor CD41 and c-kit expression to confirm partial differentiation of ESC into macrophage progenitors. CD11b and F4/80 were then monitored by flow cytometry, demonstrating efficient differentiation into macrophages.

(B) Substitution mutations introduced into the mouse *Il12b* NF-κB1, NF-κB2, and NF-κB3 sequences are shown at the bottom, with the WT *Il12b* promoter alignment in mouse and human shown for comparison. The NF-κB4 motif was not examined in this analysis. Substitution mutations in a downstream NFAT motif and a non-conserved motif were examined as controls.

Figure S4-4. Intrinsic Binding of NF-κB Dimers to the *Il12b* NF-κB3 and NF-κB4 Motifs

(A) EMSAs were performed with recombinant NF-κB RHR dimers expressed in E. coli. Binding was compared with increasing concentrations of c-Rel, RelA, and RelA(C46) homodimers to radiolabeled oligonucleotide problems containing the *Il12b* NF-κB3 and NF-κB4 motifs. Comparable binding of c-Rel and RelA homodimers to a probe containing a consensus NF-κB motif was examined as a control (right). Coomassie stains of the purified proteins are shown at the bottom.

(B) EMSAs were performed to examine relative binding strength of various NF-κB RHR homodimers and heterodimers to a radiolabeled probe containing the *Il12b* NF-κB3 sequence. For heterodimers, two NF-κB family members were co-expressed in E. coli. Coomassie stains of the purified are shown at the bottom.

(C) EMSAs were performed to examine relative binding strength of various NF-κB RHR homodimers and heterodimers to a radiolabeled probe containing the *Il12b* NF-κB4 sequence. For heterodimers, two NF-κB family members were co-expressed in E. coli. Coomassie stains of the purified are shown at the bottom. Binding of c-Rel and RelA homodimers to a probe containing a consensus NF-κB motif is shown at the right as a control.

Figure S4-5. Analysis of c-Rel Binding to *Il12b* Promoter Probes containing Tandem Motifs

(A) EMSAs were used to examine c-Rel RHR binding to probes containing substitution mutations in the *Il12b* NF-κB1, 2, and 3 motifs in a radiolabeled probe containing these three motifs in tandem, as in the WT *Il12b* promoter.

(B) An EMSA was used to examine binding of the c-Rel RHR to a radiolabeled probe containing all four tandem *Il12b* NF-κB motifs. A probe containing the *Il12b* NF-κB1 to NF-κB3 motifs is on the left. A probe containing the *Il12b* NF-κB1 to NF-κB4 motifs is on the right.

Figure S4-6. Analysis of Cooperative Binding of c-Rel to the *Il12b* NF-κB Motifs

(A) To examine the possibility of cooperative binding, EMSAs were performed with the c-Rel RHR and probes containing different pairs of *Il12b* NF-κB motifs at their native spacing. For these experiments, increasing concentrations of c-Rel were examined. The percentage of probe that assembles into a complex with the two sites simultaneously was monitored and compared to parallel experiments using probes containing each of the two motifs alone (not shown). (An examination of the complex containing only one bound dimer in the images shown provides a separate measure of binding strength to the highest affinity motif of the two motifs present in each probe.)

(B) Quantitative analysis of the data in panel A allows for the calculation of an approximate cooperativity index. For these calculations, the concentration of protein needed to occupy the lower affinity site on 50% of the probe molecules containing that site alone was divided by the concentration needed to occupy both sites simultaneously in the probe containing both sites. The results show that the presence of the *Il12b* NF-κB2 sequence allows greatly enhanced binding to the lower affinity NF-κB3 sequence. However, compelling cooperativity was not observed with the other motif pairs.

Figure 4-1. NF-κB Phylogenetic Analysis and Highly Selective c-Rel Requirement

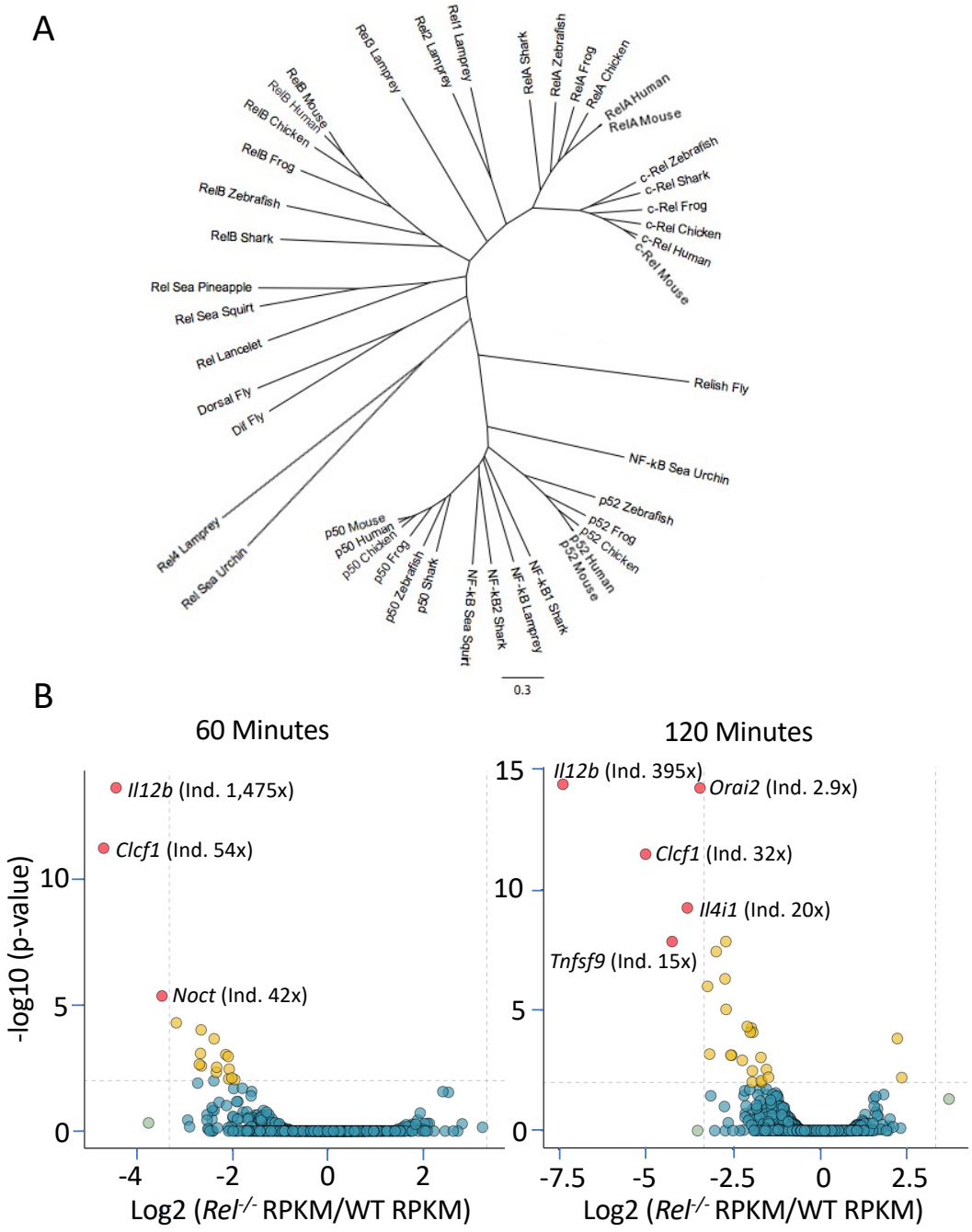


Figure 4-2. PBM and SPR Evidence of NF-κB Binding to Novel Motif Sequences

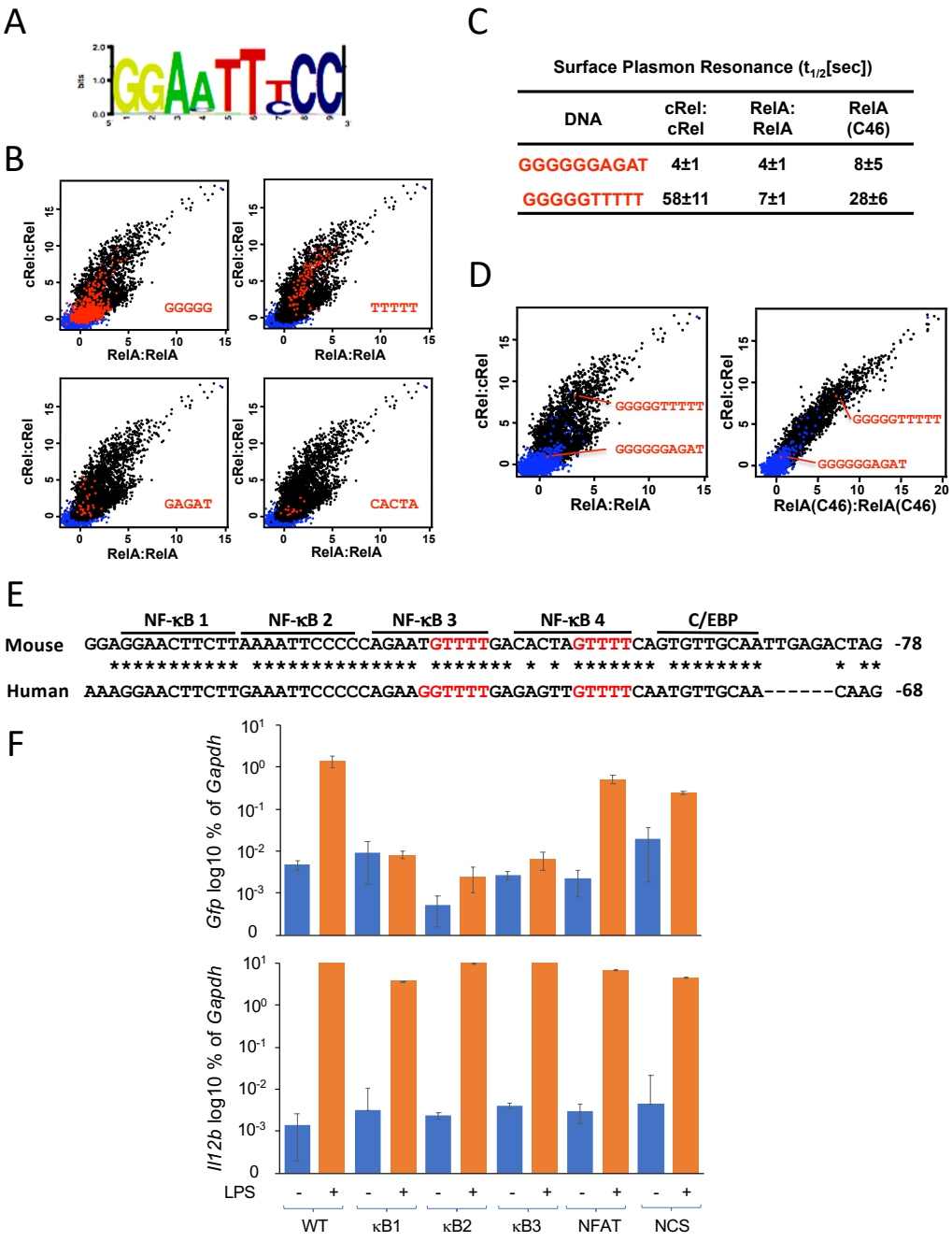
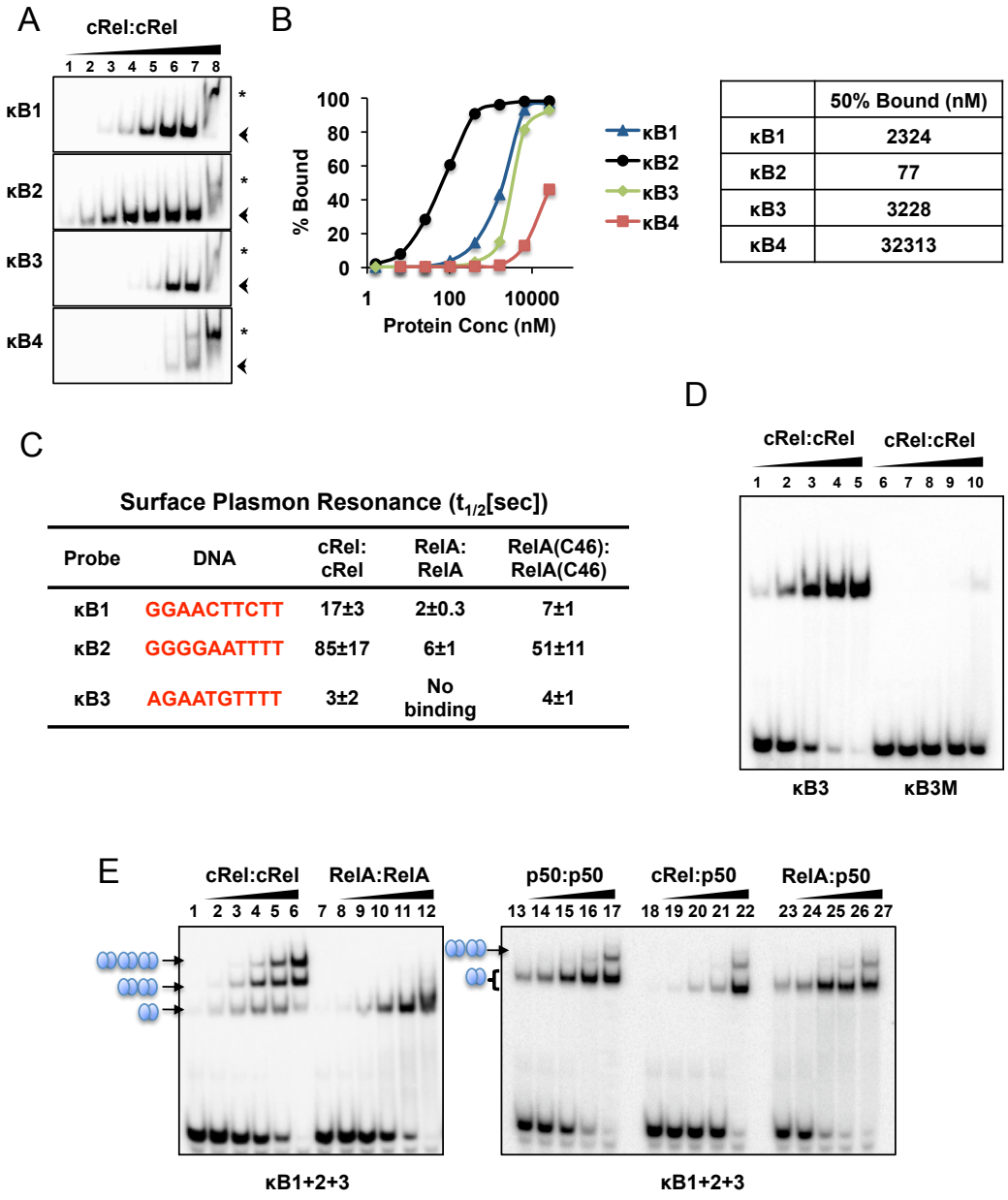


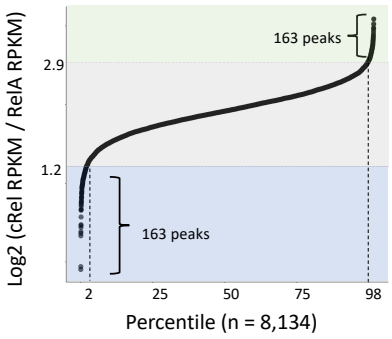
Figure 4-3. c-Rel Homodimer-Specific Binding to the *Il12b* Promoter Motifs



	50% Bound (nM)
κB1	2324
κB2	77
κB3	3228
κB4	32313

Figure 4-4. Selective Binding of c-Rel and RelA in Mouse BMDMs Examined by ChIP-seq

A

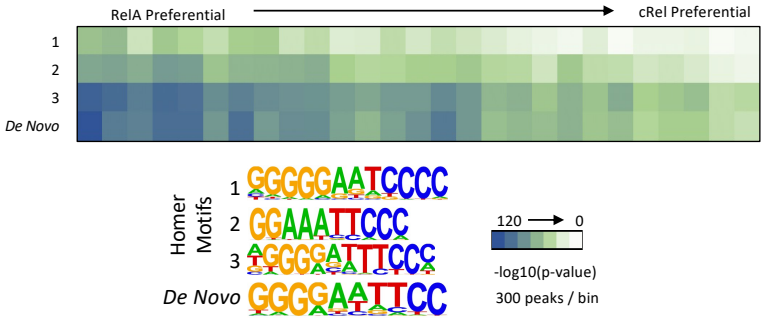


Criteria	Peak Number
> 3	113
3 – 1.3	7,549
1.3 – 0.6	467
0.6 – 0.3	5
< 0.3	0

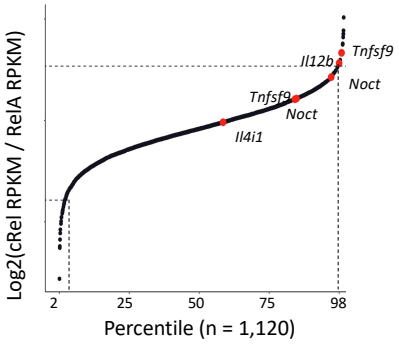
B

De Novo Motifs	p-value	% of Targets	% of Background	cRel/ RelA	Percent (300 peaks/ analysis)
	1e-46	61.67%	22.62%	> 2.7	Top 4%
	1e-84	38.46%	3.51%	~1.9	Middle 4%
	1e-117	47.33%	3.66%	< 1.25	Bottom 4%

C



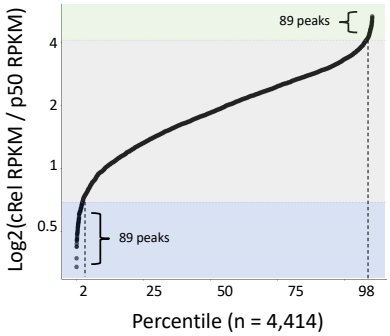
D



Gene	Peak Rank	Percentile	FC
Tnfsf9	8	0.7%	3.2
Il12b	14	1.3%	2.9
Noct	29	2.5%	2.7
Tnfsf9	124	11.0%	2.3
Noct	128	11.4%	2.3
Il4i1	351	31.3%	1.9

Figure 4-5. Selective Binding of c-Rel and p50 in Mouse BMDMs Examined by ChIP-seq

A



Criteria	Peak Number
> 3	2,646
3 – 1.3	1,078
1.3 – 0.6	492
0.6 – 0.3	161
< 0.3	37

B

Rank	De Novo Motifs	p-value	% of Targets	% of Background	cRel/ p50	Percent (300 peaks/ analysis)
1		1e-145	49.3%	2.675	>12	Top 7%
2		1e-41	46.33%	13.8%		
1		1e-108	49.16%	4.76%	~3.5	Middle 7%
3		1e-31	41.81%	4.33%		
1		1e-393	75.33%	0.81%	< 0.83	Bottom 7%
2		1e-24	8.0%	0.33%		

C

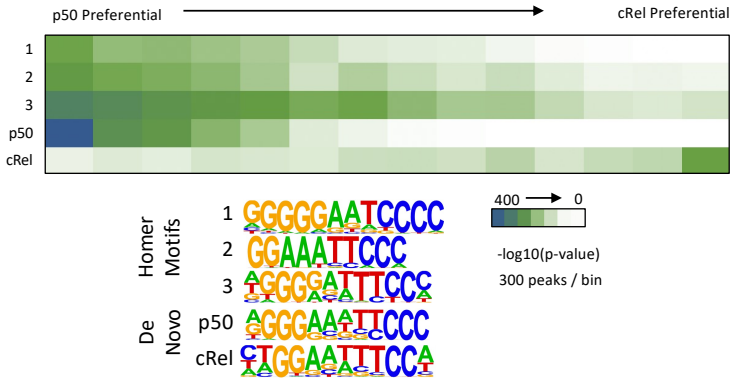


Figure 4-6. Late Evolution of the c-Rel and RelA DNA-Binding Differences

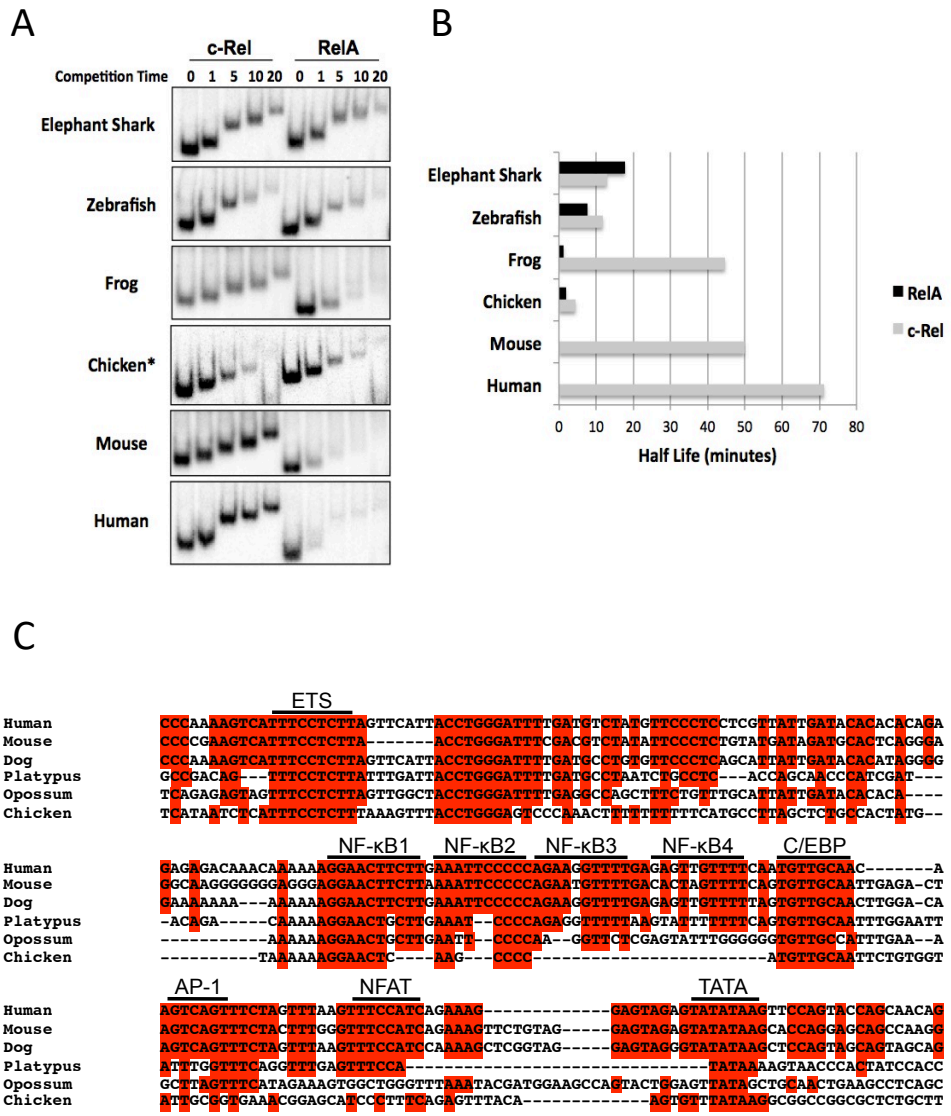


Figure S4-1. Lamprey NF-κB RHR Sequences

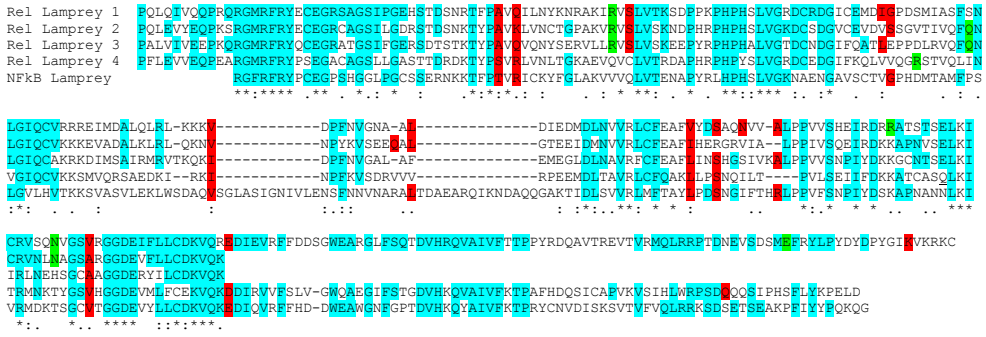


Figure S4-2. Functional Analysis of Non-Consensus NF-κB Motifs in the Mouse *Il12b* Promoter by c-Rel and RelA Overexpression in HEK 293T Cells

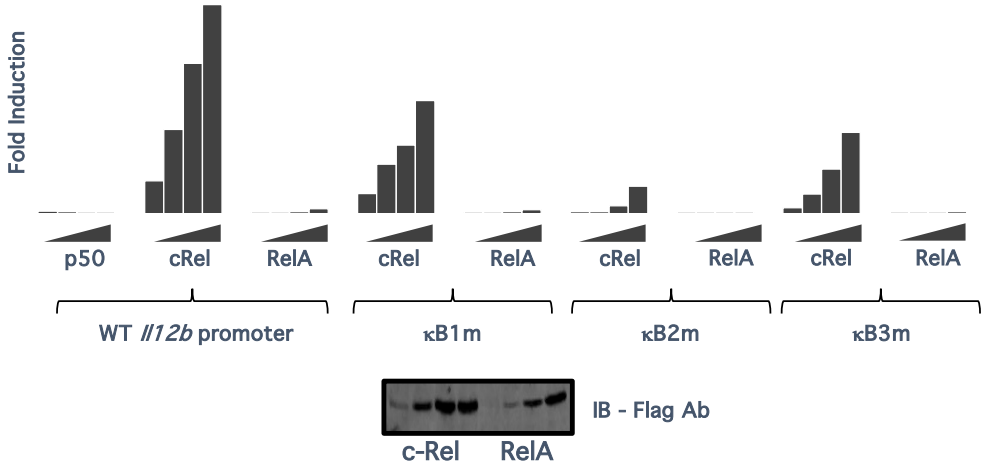
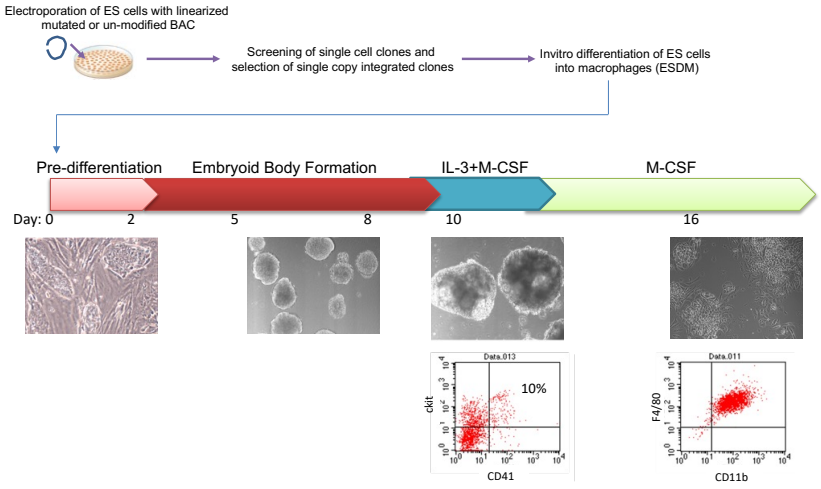


Figure S4-3. Functional Analysis of Non-Consensus NF-κB Motifs in the Mouse *Il12b* Promoter Using Recombinant BACs in ESC-Derived Macrophages

A



B

Figure S4-4. Intrinsic Binding of NF-κB Dimers to the *Il12b* NF-κB3 and NF-κB4 Motifs

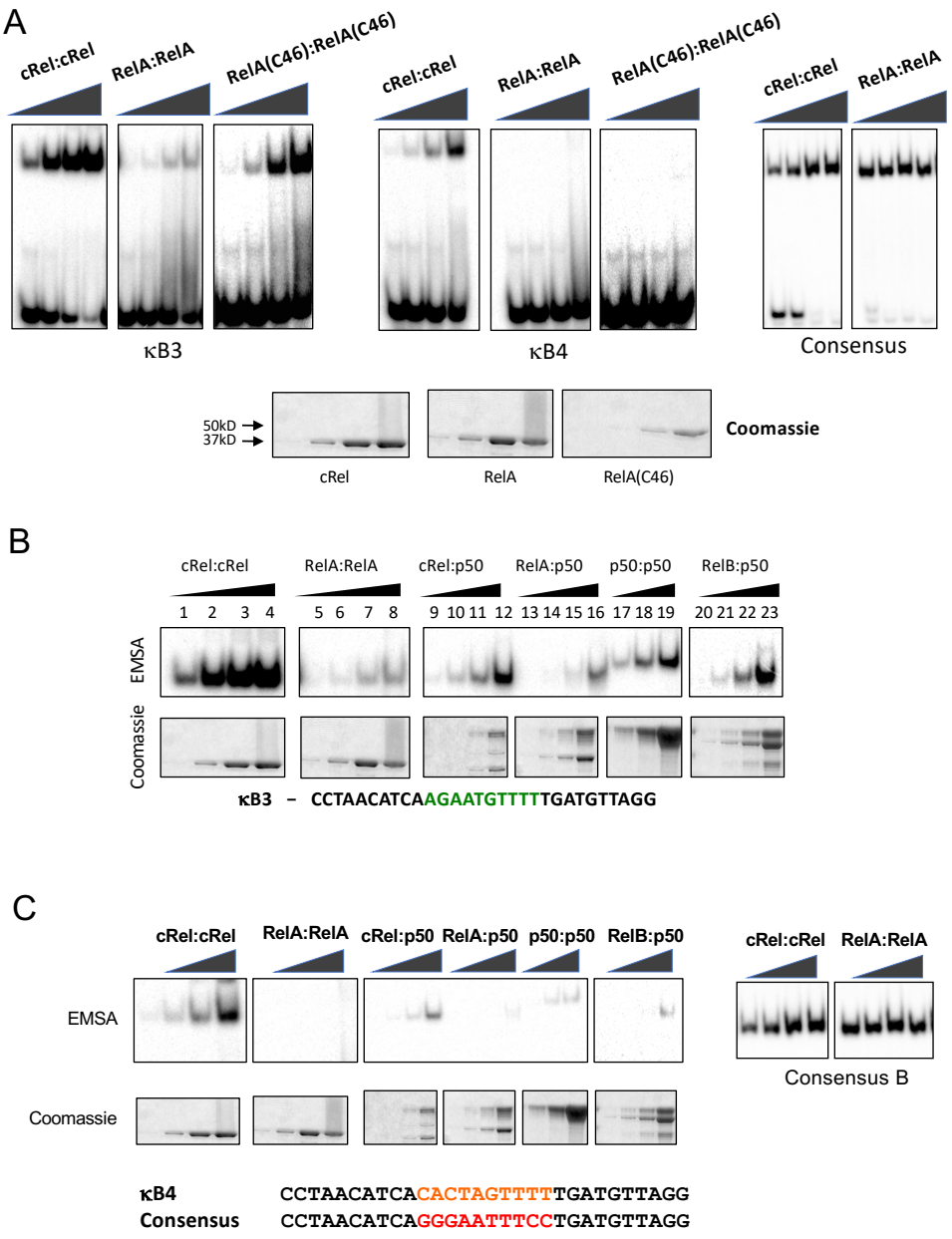


Figure S4-5. Analysis of c-Rel Binding to *Il12b* Promoter Probes containing Tandem Motifs

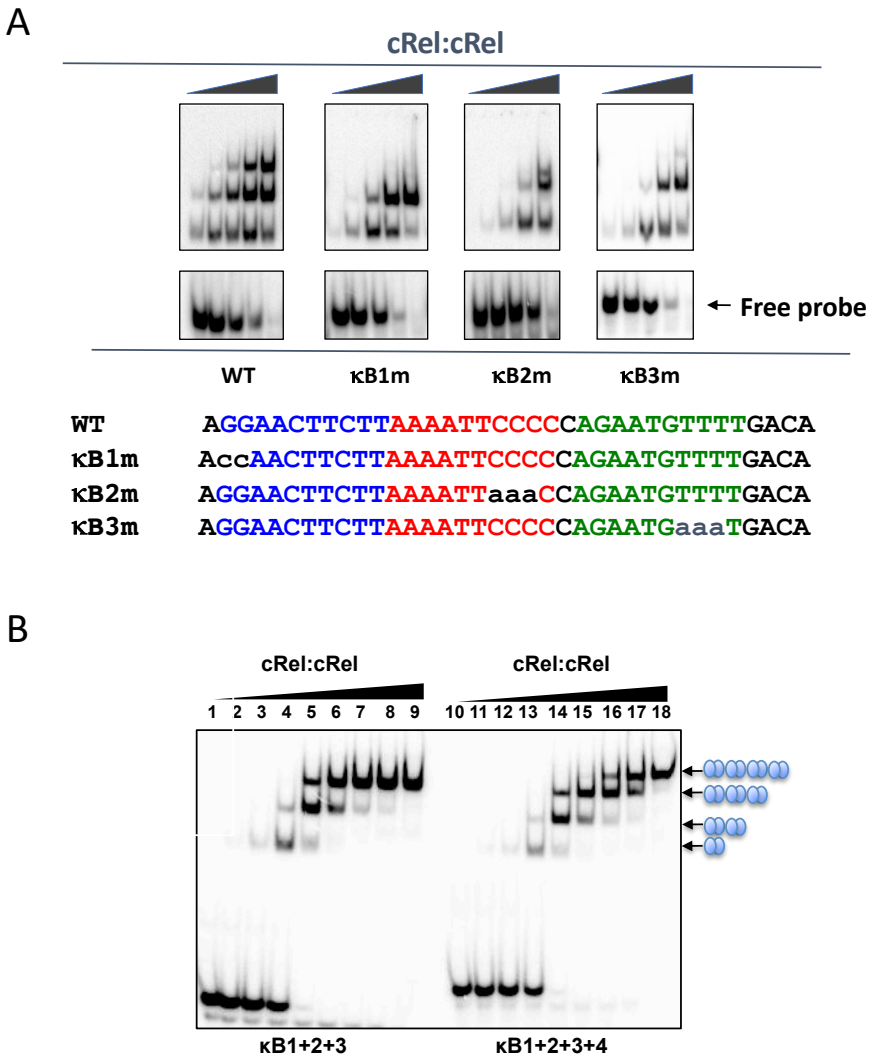
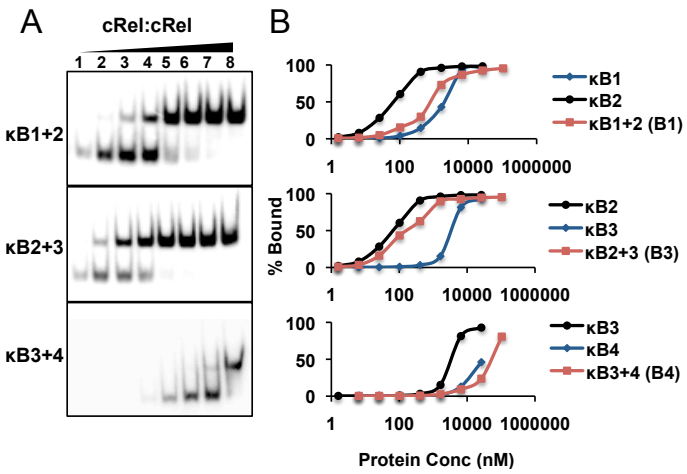


Figure S4-6. Analysis of Cooperative Binding of c-Rel to the *Il12b* NF-κB Motifs



C

Relative Cooperativity

$\frac{\kappa B1}{\kappa B1_{(\kappa B1+2)}}$	2.4
$\frac{\kappa B3}{\kappa B3_{(\kappa B2+3)}}$	15.5
$\frac{\kappa B4}{\kappa B4_{(\kappa B3+4)}}$	0.5

WORKS CITED

Andersen, K.G., Nissen, J.K, and Betz, A.G. 2012. Comparative genomics reveals key gain-of-function events in Foxp3 during regulatory T cell evolution. Front. Immunol. 3, 113.

Andersson, D.I., Hultqvist, J.J., and Näsvall, J. 2015. Evolution of new functions de novo and from preexisting genes. Cold Spring Harb. Perspect. Biol. 7, a017996.

Barish, G.D., Yu, R.T., Karunasiri, M., Ocampo, C.B., Dixon, J., Benner, C., Dent, A.L., Tangirala, R.K., and Evans, R.M. (2010). Bcl-6 and NF-κB cistromes mediate opposing regulation of the innate immune response. Genes Dev. *24*, 2760–2765.

Beg, A.A., Sha, W.C., Bronson, R.T. Ghosh, S., and Baltimore, D. 1995. Embryonic lethality and liver degeneration in mice lacking the RelA component of NF-kappa B. Nature 376, 167-170.

Bhatt, D.M, Pandya-Jones, A., Tong, A.J., Lissner, M., Barozzi, I., Natoli, G., Black, D.L., and Smale, S.T. 2012. Transcript dynamics of pro-inflammatory genes uncovered by sequence analysis of subcellular RNA fractions. Cell 250, 279-290.

Boehm, T., Hirano, M., Holland, S.J., Das, S., Schorpp, M., and Cooper, M.D. 2018. Evolution of alternative adaptive immune systems in vertebrates. Annu. Rev. Immunol. 36, 19-42.

Carmona, L.M., and Schatz, D.G. 2017. New insights into the evolutionary origins of the recombination-activating gene proteins and V(D)J recombination. FEBS J. 284, 1590-1605.

Carroll, S.B. 2005. Evolution at two levels: On genes and form. PLoS Biol. 3, e245.

Cheatle Jarvela, A.M., and Hinman, V.F. 2015. Evolution of transcription factor function as a mechanism for changing metazoan developmental gene regulatory networks. EvoDevo 6, 3.

Chen, F.E., and Ghosh, G. 1999. Regulation of DNA binding by Rel/NF-κB transcription factors: structural views. Oncogene 18, 6845-6852.

Chen, Y.Q., Ghosh, S., and Ghosh, G. 1998. A novel DNA recognition mode by the NF-kappaB p65 homodimer. Nat. Struct. Biol. 5, 67-73.

Conant, G.C., and Wolfe, K.H. 2008. Turning a hobby into a job: How duplicated genes find new functions. Nat. Rev. Genet. 9, 938-950.

DeLaney, A.A., Berry, C.T., Christian, D.A., Hart, A., Bjanes, E., Wynosky-Dolfi, M.A., Li, X., Tummers, B., Udalova, I.A., Chen, Y.H., et al. 2019. Caspase-8 promotes c-Rel-dependent inflammatory cytokine expression and resistance against *Toxoplasma gondii*. Proc. Natl. Acad. Sci. USA 116, 11926-11935.

Dong, J., Jimi, E., Zhong, H., Hayden, M.S. and Ghosh, S. 2008. Repression of gene expression by unphosphorylated NF-kappaB p65 through epigenetic mechanisms. Genes Dev. 22, 1159-1173.

Gera, T., Jonas, F., More, R., and Barkai, N. 2022. Evolution of binding preferences among whole-genome duplicated transcription factors. eLife 11, e73225.

Gerondakis, S., Grossman, M., Nakamura, Y, Pohl, T., and Grumont, R. 1999. Genetic approaches in mice to understand Rel/NF-kappaB and IkappaB function: transgenics and knockouts. Oncogene 18, 6888-6895.

Gerondakis, S., and Siebenlist, U. 2010. Roles of the NF-κB pathway in lymphocyte development and function. Cold Spring Harb. Perpsect. Biol. 2, a000182.

Ghosh, G., van Duyne, G., Ghosh, S., and Sigler, P.B. 1995. Structure of NF-kappaB p50 homodimer bound to a kappa B site. Nature 373, 303-310.

Gilmore, T.D., and Gerondakis, S. 2011. The c-Rel transcription factor in development and disease. Genes Cancer 2, 695-711.

Grumont, R., Hochrein, H., O'Keeffe, M., Gugasyan, R., White, C., Caminschi, I., Cook, W., and Gerondakis, S. 2001. c-Rel regulates interleukin 12 p70 expression in CD8(+) dendritic cells by specifically inducing p35 gene transcription. J. Exp. Med. 194, 1021-1032.

Hayden, M.S., and Ghosh, S. 2012. NF-κB, the first quarter-century: remarkable progress and outstanding questions. Genes Dev. 26, 203-234.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hoffmann, A., Natoli, G., and Ghosh, G. 2006. Transcriptional regulation via the NF-κB signaling module. Oncogene 25, 6706-6716.

Holland, L.Z., and Ocampo Daza, D. 2018. A new look at an old question: when did the second whole genome duplication occur in vertebrate evolution? Genome Biology 19, 209.

Huang, D.B, Chen, Y.Q., Ruetsche, M., Phelps, C.B., and Ghosh, G. 2001. X-ray crystal structure of proto-oncogene product c-Rel bound to the CD28 response element of IL-2. Structure 9, 669-678.

Isomura, I., Palmer, S., Grumont, R.J., Bunting, K., Hoyne, G., Wilkinson, N., Banerjee, A., Proietto, A., Guyasyan, R., Wu, L., et al. 2009. c-Rel is required for the development of thymic Foxp3+ CD4 regulatory T cells. J. Exp. Med. 206, 3001-3014.

Kane, L.P., Lin, J., and Weiss, A. 2002. It's all Rel-ative: NF-kappaB and CD28 costimulation of T-cell activation. Trends Immunol. 23, 413-420.

Keller, G.M., Webb, S., and Kennedy, M. 2002. Hematopoietic development of ES cells in culture. Methods Mol. Med. 63, 209-230.

Köntgen, F., Grumont, R.J., Strasser, A., Metcalf, D., Li, R., Tarlinton, D., Gerondakis, S. 1995. Mice lacking the c-rel proto-oncogene exhibit defects in lymphocyte proliferation, humoral immunity, and interleukin-2 expression. Genes Dev. 9, 1965-1977.

Kunsch, C., Ruben, S.M., and Rosen, C.A. 1992. Selection of optimal NF-κB/Rel DNA-binding motifs, Interaction of both subunits of NF-κB with DNA is required for transcriptional activation. Mol. Cell Biol. 12, 4412,4421.

Lévy, R., Langlais, D., Béziat, V., Rapaport, F., Rao, G., Lazarov, T., Bourgey, M., Zhou, Y.J., Briand, C., Moriya, K., et al. 2021. Inherited human c-Rel deficiency disrupts myeloid and lymphoid immunity to multiple infectious agents. J. Clin. Invest. 131, e150143.

Lynch V.J., and Wagner, G.P. 2008. Resurrecting the role of transcription factor change in developmental evolution. Evolution 62-9, 2131-2154.

Mason, N., Aliberti, J., Caamano, J.C., Liou, H.C., and Hunter, C.A. 2002. Cutting Edge: Identification of c-Rel-dependent and -independent pathways of IL-12 production during infectious and inflammatory stimuli. J. Immunol. 168, 2590-2594.

Murphy, T.L., Cleveland, M.G., Kulesza, P., Magram, J., and Murphy, K.M. 1995. Regulation of interleukin 12 p40 expression through an NF-κB half-site. Mol. Cell Biol. 15, 5258-5267.

Oh, H., Grinberg-Bleyer, Y., Liao, W., Maloney, D., Wang, P., Wu, Z., Wang, J., Bhatt, D.M., Heise, N., Schmid, R.M., et al. 2017. An NF-κB transcription-factor-dependent lineage-specific transcriptional program promotes regulatory T cell identity and function. Immunity 47, 450-465.

Plevy, S.E., Gemberling, J.H., Hsu, S., Dorner, A.J., and Smale, S.T. 1997. 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. 17, 4572-4588.

Quinlan, A.R., and Hall, I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841-842.

Rackaityte, E., and Halkias, J. 2020. Mechanisms of fetal T cell tolerance and immune regulation. Front. Immunol. 11, 588.

Ruan, Q., Kameswaran, V., Zhang, Y., Zheng, S., Sun, J., Wang, J., DeVirgiliis, J., Liou, H.C., Beg, A.A., and Chen, Y.H. 2011. The Th17 immune response is controlled by the Rel-RORγ-RORγT transcriptional axis. J. Exp. Med. 208, 2321-2333.

Samstein, R.M., Josefowicz, S.Z., Arvey, A., Treuting, P.M., and Rudensky, A.Y. 2012. Extrathymic generation of regulatory T cells in placental mammals mitigates maternal-fetal conflict. Cell 150, 29-38.

Sanjabi, S., Hoffmann, A., Liou, H.C., Baltimore, D., and Smale, S.T. 2000. Selective requirement for c-Rel during interleukin-12 p40 induction in macrophages. Proc. Natl. Acad. Sci. USA 97, 12705-12710.

Sanjabi, S., Williams, K.J., Saccani, S., Zhou, L., Hoffmann, A., Gerondakis, S., Natoli, G., and Smale, S.T. 2005. A c-Rel subdomain responsible for enhanced DNA-binding affinity and selective gene activation. Genes Dev. 19, 2138-2151.

Sen, R., and Baltimore, D. 1986. Inducibility of kappa immunoglobulin enhancer-binding protein Nf-kappa B by a posttranslational mechanism. Cell 47, 921-928.

Siggers, T., Chang, A.B., Teixeira, A., Wong, D., Williams, K.J., Ahmed, B., Ragoussis, J., Udalova, I.A., Smale, S.T., and Bulyk, M.L. 2011. Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-κB family DNA binding. Nat. Immunol. 13, 95-102.

Shapiro, V.S., Truitt, K.E., Imboden, J.B., and Weiss, A. 1997. CD28 mediates transcriptional upregulation of the interleukin-2 (IL-2) promoter through a composite element containing the CD28RE and NF-IL-2B AP-1 sites. Mol. Cell Biol. 17, 4051-2058.

Smale, S.T. 2011. Hierarchies of NF-κB target-gene regulation. Nat. Immunol. 12, 689-694.

Smale, S.T. 2012. Dimer-specific regulatory mechanisms within the NF-κB family of transcription factors. Immunological Rev. 246, 193-204.

Smith, J.J., Timoshevskaya, N., Ye, C., Holt, C., Keinath, M.C., Parker, H.J., Cook, M.E., Hess, J.E., Narum, S.R., Lamanna, F., et al. 2018. The sea lamprey germline genome provides insights into programmed genome rearrangements and vertebrate evolution. Nat. Genet. 50, 270-277.

Tong, A.J., Liu, X., Thomas, B.J., Lissner, M.M., Baker, M.R., Senagolage, M.D., Allred, A.A., Barish, G.D., and Smale, S.T. 2016. A stringent systems approach uncovers gene-specific mechanisms regulating inflammation. Cell 165, 165-179.

Tumang, J.R., Owyang, A., Andjelic, S., Jin Z., Hardy, R.R., Liou, M.L., and Liou, H.C. 1998. c-Rel is essential for B lymphocyte survival and cell cycle progression. Eur. J. Immunol. 28, 4299-4312.

Tousto, L. 2011. NF-κB family of transcription factors: biochemical players of CD28 co-stimulation. Immunol. Lett. 135, 1-9.

Venkatesh, B., Lee, A.P., Ravi, V., Maurya, A.K., Lian, M.M., Swann, J.B., Ohta, Y., Flajnik, M.F., Sutoh, Y., Kasahara, M., et al. 2014. Elephant shark genome provides unique insights into gnathostome evolution. Nature 505, 174-179.

Wan, F., Anderson, D.E., Barnitz, R.A., Snow, A., Bidere, N., Zheng, L., Hedge, V., Lam, L.T., Staudt, L.M., Levens, D., et al. 2007. Ribosomal protein S3: a KH domain subunit in NF-kappaB complexes that mediates selective gene regulation. Cell 131, 927-939.

Wang, J., Wang, X, Hussain, S., Zheng, Y, Sanjabi, S., Ouaaz, F., and Beg, A.A. 2007. Distinct roles of different NF-κB subunits in regulating inflammatory and T cell stimulatory gene expression in dendritic cells. J. Immunol. 178, 6777-6788.

Warming, S., Costantino, N., Court, D.L., Jenkins, N.A., and Copeland, N.G. 2005. Simple and highly efficient BAC recombineering using *galK* selection. Nucleic Acids Res. 33, e36.

Williams, L.M., and Gilmore, T.D. 2020. Looking down on NF-κB. Molec. Cell. Biol. 40, e00104-20.

Zhang, H., Bi, J., Yi, H., Fan, T., Ruan, Q., Cai, L, Chen, Y.H., and Wan, X. 2017. Silencing of c-Rel in macrophages dampens Th1 and Th17 immune responses and alleviates experimental autoimmune encephalomyelitis in mice. Immunol. Cell Biol. 95, 593-600.

Zheng, Y., Josefowicz, S., Chaudhry, A., Peng, X.P., Forbush, K., and Rudensky, A.Y. 2010. Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate. Nature 463, 808-812.

Zhong, H., Voll, R.E., and Ghosh, S. 1998. Phosphorylation of NF-κB p65 by PKA stimulates transcriptional activity by promoting a novel bivalent interaction with the coactivator CBP/p300. Mol. Cell 1, 661-671.

CHAPTER 5

Concluding Remarks

In the immune system, proper gene expression is crucial for the resolution of pathogenic infection. Activation of gene expression is intricately regulated through transcriptional, post-transcriptional, translational, and post-translational mechanisms[128–130]. Dysregulation of proinflammatory genes can result in the aberrant activation of immune cells, leading to chronic inflammation, autoimmune disease, and dysbiosis. Given that transcriptional regulation is the primary means of gene regulation, our study focused on this mechanism.

Sequence-specific transcription factors (TFs) provide a rapid method for cells to modulate gene expression. TFs are often transcribed and translated in unstimulated cells, which poises them for activation. Upon stimulation, TFs are quickly activated through post-translational modifications of the TF itself or of a negative regulator. Various TFs are activated in immune cells following an infection of the host organism. However, the NF-κB family is one of the master regulators of immune cell activation and thus, the focus of this research. Despite extensive studies exploring the role of NF-κB since its discovery almost 40 years ago, numerous questions remain regarding NF-κB selective functions.

Transcription factors, like NF-κB, are grouped into families based on structural and functional properties. The NF-κB family contains five subunits that can combine into 15 dimers. Some TF families, like the NF-κB family, exhibit a high degree of combinatorial complexity. Often different members of a TF family are expressed in the same cells and are activated through conserved mechanisms. Thus, the question remains: what are the specific roles of individual members of a TF family in response to inflammatory stimuli?

This study employed the use of bone marrow-derived macrophages (BMDMs) to explore the combinatorial complexity of the NF-κB family with unprecedented depth. We used LPS to mimic the stimulation of cells by gram-negative bacteria. The LPS:TLR4 pathway has been well

classified as a potent activator of NF-κB. With RNA-seq from transgenic or C47BL/6 BMDMs stimulated with LPS, we uncovered unique roles for individual subunits of NF-κB in the inflammatory immune response. With ChIP-seq, ATAC-seq, Sequential ChIP-seq, EMSAs, and co-immunoprecipitations, we defined novel mechanisms underlying dimer-specific functions of NF-κB.

In Chapter 2, we began by comprehensively studying and optimizing the capture of protein:DNA interactions. The advent of genomics approaches such as ChIP-seq and CUT&TAG, creates an opportunity to study transcriptional regulation in an unbiased manner. Prior to exploring immunological questions of NF-κB selective function, we systematically explored the effect of crosslinking on ChIP-seq. We found that two chemical crosslinkers commonly used during ChIP-seq dramatically affected the results. Higher concentrations of crosslinkers increased the total number of binding sites but decreased the specificity of interactions. The highest crosslinking conditions we explored led to ChIP-seq results with an abundance of aberrant interactions, which did not contain appropriate motifs. After thoroughly optimizing ChIP-seq conditions, we investigated the ability of ChIP-seq to capture NF-κB selective interactions. We performed the first high-quality quantitative analysis of NF-κB subunits in the cell context and found that a limited number of sites in the genome exhibited selective binding. Interestingly, we revealed unique NF-κB motifs that specify preferential binding sites. This was consistent with previous biochemical and structural studies. After understanding the proper conditions for ChIP-seq, we compared ChIP-seq to CUT&TAG (an assay that circumvents the necessity of crosslinking altogether). We found that while these techniques are reproducible, both assays captured off-target interactions.

In Chapter 3, we explored the role of p50 in macrophage activation in response to bacterial stimulation. Comparing RNA-seq from *Nfkb1*$^{-/-}$ and wild-type BMDMs, we defined the non-redundant role that p50 plays in macrophage activation. We discovered that p50 is necessary for

the activation of a surprisingly small group of inflammatory genes. Given that p50 is known to form the most abundant dimer of NF-κB with RelA, which regulates hundreds of genes, it was surprising that p50 is only required for the expression of about 30 genes. After defining the role of p50 in macrophage activation, we began exploring the underlying mechanism of dimer-specific regulation. Using low crosslinking conditions, we elucidated p50-preferentially bound sites and revealed that they are not enriched near p50-dependent genes. This suggested that p50 homodimers do not play a direct role in the modulation of p50 dependent genes. While p50 homodimers have been implicated in diverse mechanisms of gene regulation, we were unable to re-capitulate these roles in LPS-activated macrophages. It is possible that p50 homodimers have functional roles in other cell types or in response to different stimuli. However, future studies will need to perform in-depth analyses to further describe these putative roles for p50 homodimers in transcriptional regulation.

When comparing gene dependencies between p50 and nuclear IκBs, we observed a striking overlap in gene dependencies between p50 and IκBζ. Importantly, p50 and IκBζ selectively regulate a group of key inflammatory genes. These data suggest there is differential regulation of either p50 or IκBζ, which would allow for fine-tunned immunomodulation. Our data supported that IκBζ is differentially regulated by TNF. While there are hundreds of genes with differential expression between TNF or LPS-activated macrophages, we found that the p50:IκBζ pathway can explain a significant number of these differentially expressed genes. Moreover, we distinguished functional binding sites by defining IκBζ binding sites with dependence on p50.

Our analyses remain inconclusive in fully explaining the reasons underlying the dependence of genes on p50 and IκBζ. For hundreds of genes, p50:RelA heterodimers are sufficient for transcriptional activation. The essential role of IκBζ at cis-regulatory elements

(where p50:RelA heterodimers are also present) is still unclear. One possible explanation is that IκBζ associates with a necessary co-activator, or the 3D chromatin structure requires further stabilization of NF-κB, a role that IκBζ may fulfill. However, future studies are needed to fully elucidate the role of IκBζ in transcriptional regulation.

In Chapter 4, we explored another interesting instance of NF-κB dimer-specific function by investigating the role of c-Rel in macrophage activation. Our studies revealed that in macrophages stimulated with LPS, c-Rel selectively regulates *Il12b,* a cytokine needed for proper T-cell differentiation. We illustrated the unique ability of c-Rel homodimers to bind with higher affinity to non-consensus NF-κB motifs through structural, biochemical, and genomics approaches. The use of ChIP-seq with low crosslinking enabled us to capture c-Rel preferential interactions and reveal a non-consensus motif that specifies these binding sites *in vivo*. This select role of c-Rel is consistent with the evolutionary divergence of NF-κB proteins to regulate the interplay between innate and adaptive immunity. These data suggest there may be a differential expression of c-Rel in certain biological contexts. Thorough quantitative studies exploring the expression of c-Rel in response to a diverse array of agonists could reveal a condition where c-Rel and RelA exhibit differential expression. Given that c-Rel appears to play a unique role in modulating genes that connect innate and adaptive immunity, investigating the role of c-Rel in adaptive immune cell activation would also be intriguing. Furthermore, while the divergence of c-Rel and RelA aligns with the evolution of the adaptive immune system, the underlying factors that drove this divergence remain unclear.

A critical aspect of mechanistically understanding transcriptional regulation is accurately annotating functional enhancers. Due to the looping structure of DNA, enhancers that are up to a mega-base away can influence gene expression. In the studies presented in this thesis, we explored the enrichment of binding sites near genes with strong induction or dependence to

identify functional enhancers. The assumption is that strong enrichment of a group of enhancers with a unique property (such as a distinct motif or binding dependence) near dependent genes would not occur randomly, indicating that enriched enhancer sites are likely to be transcriptionally relevant.

To confirm the function of an enhancer in modulating gene expression, mutations of the binding site with CRISPR/Cas9 homology-directed repair (HDR) are needed. While CRISPR/Cas9 HDR is a powerful tool, generating a pure colony is time-consuming and requires an immortalized cell to undergo single-cell colony expansion. CRISPR is now being adapted using deactivated-Cas9 to modulate gene expression with techniques like CRISPRi/a. These techniques create opportunities to study enhancer function by recruiting factors artificially and combinatorically to cis-regulatory regions to assess the effect on transcription.

We have successfully elucidated select roles for individual NF-κB dimers through in-depth analyses. However, our overall understanding of selective transcriptional regulation is just beginning. New technologies will enable future studies to continue expanding our knowledge of selective transcriptional regulation. Understanding how TFs from the same family preferentially activate subsets of genes is critical to our knowledge of gene regulation. Moreover, a mechanistic understanding of pro-inflammatory gene regulation will generate a clearer view of the processes driving inflammation and expose new therapeutic targets.

WORKS CITED

1. Simen Zhao, B., Roundtree, I. A. & He, C. S-Adenosyl methionine post-transcriptional gene regulation by mRNA modifications. Nat. Publ. Gr. 18, 31 (2016).
2. Cramer, P. Organization and regulation of gene transcription. Nature doi:10.1038/s41586-019-1517-4 (2019).
3. Orphanides, G. & Reinberg, D. A unified theory of gene expression. Cell 108, 439–451 (2002).
4. Furman, D. et al. Chronic inflammation in the etiology of disease across the life span. Nat. Med. 2019 2512 25, 1822–1832 (2019).
5. Hayden, M. S. & Ghosh, S. Shared principles in NF-κB signaling. Cell 132, 344–362 (2008).
6. Ghosh, S. & Hayden, M. S. Celebrating 25 years of NF-κB research. Immunol. Rev. 246, 5–13 (2012).
7. Zhang, Q., Lenardo, M. J. & Baltimore, D. 30 years of NF-κB: a blossoming of relevance to human pathobiology. Cell 168, 37–57 (2017).
8. O'Dea, E. & Hoffmann, A. The regulatory logic of the NF-kappaB signaling system. Cold Spring Harb. Perspect. Biol. 2, (2010).
9. Gilmour, D. S. & Lis, J. T. Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. Proc. Natl. Acad. Sci. U. S. A. 81, 4275 (1984).
10. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-wide mapping of in vivo protein-DNA interactions. Science (80-.). 316, 1497–1502 (2007).
11. Hoon Kim, T. & Ren, B. Genome-wide analysis of protein-DNA interactions. (2006). doi:10.1146/annurev.genom.7.080505.115634

12. Kaya-Okur, H.S. et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. doi:10.1038/s41467-019-09982-5 (2019).

APPENDIX A

ARTICLES
https://doi.org/10.1038/s41590-020-0695-4

nature immunology

Interferon-mediated reprogramming of membrane cholesterol to evade bacterial toxins

Quan D. Zhou[1,2], Xun Chi[3], Min Sub Lee[1], Wei Yuan Hsieh[3], Jonathan J. Mkrtchyan[3], An-Chieh Feng[3], Cuiwen He[4], Autumn G. York[1,5,6], Viet L. Bui[3], Eliza B. Kronenberger[3], Alessandra Ferrari[7], Xu Xiao[7], Allison E. Daly[3], Elizabeth J. Tarling[4], Robert Damoiseaux[1], Philip O. Scumpia[3,4], Stephen T. Smale[3], Kevin J. Williams[8], Peter Tontonoz[7,8] and Steven J. Bensinger[1,3]✉

Plasma membranes of animal cells are enriched for cholesterol. Cholesterol-dependent cytolysins (CDCs) are pore-forming toxins secreted by bacteria that target membrane cholesterol for their effector function. Phagocytes are essential for clearance of CDC-producing bacteria; however, the mechanisms by which these cells evade the deleterious effects of CDCs are largely unknown. Here, we report that interferon (IFN) signals convey resistance to CDC-induced pores on macrophages and neutrophils. We traced IFN-mediated resistance to CDCs to the rapid modulation of a specific pool of cholesterol in the plasma membrane of macrophages without changes to total cholesterol levels. Resistance to CDC-induced pore formation requires the production of the oxysterol 25-hydroxycholesterol (25HC), inhibition of cholesterol synthesis and redistribution of cholesterol to an esterified cholesterol pool. Accordingly, blocking the ability of IFN to reprogram cholesterol metabolism abrogates cellular protection and renders mice more susceptible to CDC-induced tissue damage. These studies illuminate targeted regulation of membrane cholesterol content as a host defense strategy.

Cholesterol is the most abundant lipid in the plasma membrane and is an important determinant of the membrane's biochemical and biophysical properties[1–4]. Cellular cholesterol homeostasis is maintained through a combination of synthesis, import, storage and efflux pathways. While each pathway may be engaged independently, the pathways are often co-regulated to ensure that sufficient cholesterol is available for a cell and to avoid deleterious accumulation. Recent studies indicate that sterol metabolism pathways of immune cells, in particular macrophages, can be rapidly reprogrammed to support their effector functions[5–8]. Toll-like receptors (TLRs) and IFNs have been shown to mediate reprogramming of cholesterol homeostasis by rapidly decreasing cholesterol synthetic flux[9,10], increasing cholesterol ester storage in lipid droplets[11,12] and stimulating the production of cholesterol derivatives such as 25HC[5,13,14]. This reprogramming of sterol homeostasis facilitates host defense responses through a variety of mechanisms, including blockade of viral entry, regulation of innate immune receptor signaling, immune cell skewing and increasing macrophage phagocytic capacity. Thus, coordinated regulation of sterol metabolic pathways in immune cells is now viewed as a fundamental component of the host defense response.

Given the essential requirement for cholesterol in animal membranes, it is not surprising that microorganisms and viruses have evolved strategies that take advantage of cholesterol availability in membranes to facilitate pathogenesis[15–18]. CDCs are a group of secreted toxins produced mainly by Gram-positive bacteria that depend on membrane cholesterol for their pathogenic function[19–21]. CDC binding to cholesterol in the plasma membrane facilitate CDC oligomerization and pore formation, resulting in loss of membrane integrity, reduced efficacy of antimicrobial function and cell death[22,23]. Cells have repair mechanisms by which they minimize the deleterious effects of CDC toxins. Damaged plasma membranes can be internalized and resealed[24,25], or they may bleb and be subsequently shed as microvesicles[26]. However, it appears that phagocytes attempting membrane repair may become hyporesponsive to proinflammatory stimuli through shedding of TLRs and cytokine receptors[23]. Thus, host membrane repair mechanisms, which preserve phagocyte survival, also reduce their effectiveness for antimicrobial host defense. Whether phagocytes or other immune cells have evolved specific strategies that facilitate evasion from these pore-forming toxins remains less well understood.

Here we report that activation of pattern recognition receptor (PRR) signaling on macrophages influences susceptibility to CDC-induced loss of membrane integrity and function. We show that IFN signals convey resistance to CDC-mediated membrane damage on phagocytes by reprogramming their cellular cholesterol metabolism, whereas myeloid differentiation primary-response protein 88 (Myd88)-dependent TLR signals do not induce protection. Mechanistic studies revealed that IFNs mediate resistance to CDCs by decreasing their binding to the plasma membrane. The decrease in CDC binding requires coordinated inhibition of new cholesterol synthesis and an increase in the esterification of a small

[1]Department of Molecular and Medical Pharmacology, University of California, Los Angeles, Los Angeles, CA, USA. [2]Department of Surgical Oncology, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, P.R. China. [3]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, Los Angeles, CA, USA. [4]Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, USA. [5]Department of Immunobiology, Yale University School of Medicine, New Haven, CT, USA. [6]Howard Hughes Medical Institute, Yale University, New Haven, CT, USA. [7]Department of Pathology and Laboratory Medicine, University of California, Los Angeles, Los Angeles, CA, USA. [8]Department of Biological Chemistry, University of California, Los Angeles, Los Angeles, CA, USA. ✉e-mail: sbensinger@mednet.ucla.edu

amount of preexisting cellular cholesterol. IFN-mediated changes in cholesterol homeostasis were also dependent on the endogenous production of 25HC, and we found that loss of the enzyme cholesterol 25-hydroxylase (CH25H) resulted in increased binding of CDC proteins at the plasma membrane. Accordingly, perturbations in sterol metabolic pathways abrogated the protective effects of IFNs for phagocytes and rendered mice more susceptible to CDC-induced tissue damage. Together, these findings indicate that IFN-mediated changes to sterol metabolism convey resistance to CDCs by decreasing a small pool of cholesterol in the plasma membrane and suggest that specifically targeting this pool of cholesterol could facilitate host defense to CDC-producing pathogenic bacteria.

Results

IFNs mediate resistance to CDCs. Activation of macrophages through TLR and cytokine signals reprograms their lipid metabolism[5–8,10,27], leading us to ask whether these changes in lipid metabolism alter sensitivity to CDCs. To address this question, mouse bone marrow-derived macrophages (BMDMs) were stimulated with TLR1/TLR2, TLR3, TLR4, TLR7 or TLR9 agonist for 24 h and then challenged with the CDC perfringolysin O (PFO) for up to 60 min. The membrane-impermeable dye propidium iodide (PI) was added to the medium just before PFO treatment to assess pore formation and loss of membrane integrity but not cell death in macrophages. As expected, PFO increased the percentage of PI-positive (PI^+) cells over time in unstimulated macrophages (Fig. 1a and Extended Data Fig. 1a). Activation of macrophages with TLR1/TLR2, TLR4, TLR7 or TLR9 agonist resulted in little or no change in sensitivity to PFO treatment (Fig. 1a and Extended Data Fig. 1a,b), whereas TLR3-mediated activation markedly decreased macrophage sensitivity to PFO treatment (Fig. 1a and Extended Data Fig. 1a).

Unexpectedly, TLR3 activation, which is associated with antiviral responses, provided a substantial measure of protection to PFO-induced pore formation. Because a primary inflammatory outcome of TLR3 signaling is the generation of type I IFNs[28,29], we tested whether type I IFN stimulation alone would convey resistance to PFO. We observed that treatment of macrophages with IFN-β overnight also conveyed resistance to PFO (Fig. 1b and Extended Data Fig. 1c). IFN-γ treatment of macrophage cultures also induced resistance when cells were challenged with PFO (Fig. 1b and Extended Data Fig. 1c), indicating that protection is mediated by both type I and type II IFNs. We also tested whether activation of other PRRs that generate type I IFN responses[30] would also protect macrophages when challenged with CDCs. The stimulator of IFN genes (STING; TMEM173) is activated by cyclic dinucleotides generated by intracellular microorganisms (for example, *Listeria monocytogenes*)[31]. We found that treatment of BMDM cultures with STING ligands (2′,3′-cGAMP or c-di-GMP) for 24 h also resulted in protection from the PFO challenge (Fig. 1c and Extended Data Fig. 1d). Nucleotide-binding oligomerization domain-containing protein 2 (NOD2), an intracellular receptor stimulated by peptidoglycans produced by Gram-positive microorganisms, has also been shown to induce low levels of type I IFN during bacterial infections[32]. Consistent with this, activation of NOD2 with *N*-glycolyl-MDP resulted in modest protection against the PFO challenge (Fig. 1c and Extended Data Fig. 1d). Importantly, we observed a similar pattern of IFN-mediated resistance for macrophages challenged with the CDCs streptolysin O (SLO) or anthrolysin O (ALO; Fig. 1d,e and Extended Data Fig. 1e–g), indicating that the protection conveyed by IFNs may be a general strategy for resistance to the larger group of pore-forming CDCs.

We were surprised that TLR1/TLR2 activation did not induce resistance to CDCs given the important role of this TLR signaling module in mediating immune responses to Gram-positive bacteria. However, we found that macrophages co-stimulated with the TLR1/TLR2 agonist and IFN-β acquired some measure of resistance to PFO and SLO (Fig. 1f,g and Extended Data Fig. 1h,i), indicating that IFN signals can induce a 'resistant' phenotype regardless of the PRR engaged. We also found that IFN stimulation (IFN-γ or IFN-β; 4 h) of purified neutrophils from mouse bone marrow resulted in resistance to CDC-mediated loss of membrane integrity (Fig. 1h), demonstrating the ability of IFN to induce protection in other phagocytic, innate immune cells. Importantly, IFN stimulation largely preserved the ability of macrophages to phagocytose extracellular bacteria (that is, *Staphylococcus aureus*) when cells were challenged with PFO (Fig. 1i). Together, these data demonstrate that autocrine or paracrine IFN signals induce resistance to CDC-mediated pore formation, resulting in preservation of macrophage membrane integrity and phagocytic function.

IFNs decrease binding of CDCs to the plasma membrane. To understand whether the observed changes in CDC sensitivity could be due to altered binding of CDCs to the plasma membrane, we performed studies on IFN-stimulated macrophages with fluorescently labeled ALO domain 4 (ALO-D4), a modified CDC toxin that has markedly reduced pore-forming capacity but retains near-normal binding characteristics to the plasma membrane[33–36]. We found that the binding of ALO-D4 to IFN-β- or IFN-γ-stimulated macrophages was substantially decreased when compared to control macrophages (Fig. 2a). High-throughput imaging analysis of macrophage cultures (>3,000 macrophages for each condition) stimulated with IFN-β, IFN-γ, STING or TLR3 ligand showed a marked reduction in the average intensity of ALO-D4 binding (Fig. 2b–e). By contrast, ALO-D4 binding was maintained or modestly increased in TLR1/TLR2-stimulated macrophages (Fig. 2f,g), consistent with our data indicating that TLR1/TLR2-stimulated macrophages maintained sensitivity to CDC-induced loss of membrane integrity. A similar decrease in ALO-D4 binding in response to either IFN-α or IFN-β was observed in macrophages derived from human peripheral blood mononuclear cells (hPBMCs; Fig. 2h,i). Likewise, stimulation (6 h) of purified neutrophils from mouse bone marrow with IFNs also decreased ALO-D4 binding (Extended Data Fig. 2a,b). Time-course studies showed that a decrease in ALO-D4 binding could be observed within 2 h of IFN stimulation of BMDMs (Fig. 2j,k). As expected, the decrease in ALO-D4 signal induced by IFN-β (or PRR-induced type I IFNs) was abrogated in type I IFN receptor (IFNAR)-deficient macrophages (Extended Data Fig. 2c,d). However, overnight IFN-γ treatment of IFNAR-deficient macrophages decreased ALO-D4 binding (Fig. 2l,m), demonstrating that resistance induced by type I and type II IFN cytokines are not interdependent. In combination, these data indicate that protection from CDCs is induced quickly in response to IFNs and is mediated by reduced binding of CDCs to the plasma membrane.

IFNs decrease a small pool of membrane cholesterol targeted by CDCs. Next, we asked whether IFN signals altered the binding of CDCs to the plasma membrane by decreasing macrophage cholesterol content. To address this, we performed gas chromatography–mass spectrometry (GC–MS) to quantify total cellular cholesterol levels in IFN-stimulated BMDMs. We observed that IFNs modestly increased total cholesterol in macrophages (Fig. 3a). Likewise, we observed that STING or TLR3 activation, which induces robust type I IFN responses, also increased total cholesterol (Extended Data Fig. 3a). Thus, the protection mediated by IFNs could not be attributed to a global decrease in cholesterol content. We also performed GC–MS on purified plasma membranes from quiescent and IFN-γ-stimulated macrophages; however, no difference in cholesterol pool size was observed (Fig. 3b and Extended Data Fig. 3b). We also stained BMDMs with filipin III, a bacterial product that binds to unesterified cholesterol in membranes and forms a fluorescent complex[37]. Consistent with our mass spectrometry studies, no change in filipin III staining was observed in response

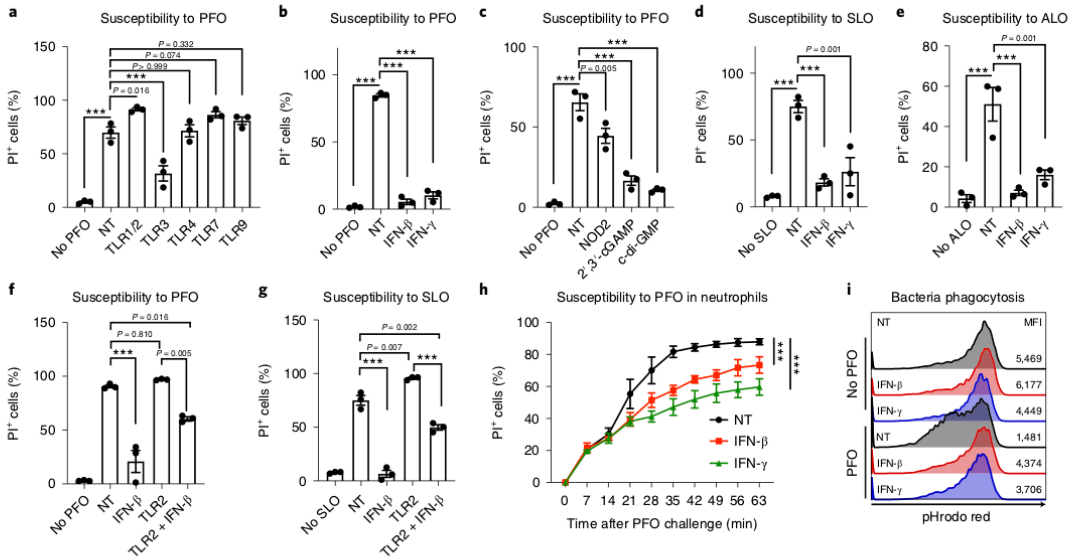


Fig. 1 | IFN signaling mediates resistance to CDCs. a, Percentage of PI+ BMDMs treated with the TLR1/TLR2 agonist (Pam3CSK4; 50 ng ml^{-1}), TLR3 agonist (poly(I:C); 1 μg ml^{-1}), TLR4 agonist (LPS; 50 ng ml^{-1}), TLR7 agonist (CL307; 100 nM), TLR9 agonist (ODN1668; 100 nM) or unstimulated (NT) for 24 h and then challenged with PFO for 60 min in the presence of PI. **b**, Percentage of PI+ BMDMs treated with IFN-β (20 ng ml^{-1}) and IFN-γ (20 ng ml^{-1}) for 24 h and then challenged with PFO for 60 min in the presence of PI. **c**, Percentage of PI+ BMDMs treated with NOD2 agonist (*N*-glycolyl-MDP; 20 μg ml^{-1}) or STING agonists (2′,3′-cGAMP and c-di-GMP; 2 μg ml^{-1}) for 24 h and then challenged with PFO for 60 min in the presence of PI. **d**, Percentage of PI+ BMDMs treated with the indicated IFNs (20 ng ml^{-1}) for 24 h and then challenged with SLO for 2 h in the presence of PI. **e**, Percentage of PI+ BMDMs treated with the indicated IFNs (20 ng ml^{-1}) for 24 h and then challenged with ALO for 2 h in the presence of PI. **f**, Percentage of PI+ BMDMs treated with TLR1/TLR2 agonist (50 ng ml^{-1}) together with IFNs (100 ng ml^{-1}) for 24 h and then challenged with PFO for 60 min in the presence of PI. **g**, Percentage of PI+ BMDMs treated with TLR1/TLR2 agonist (50 ng ml^{-1}) together with IFNs (100 ng ml^{-1}) for 24 h and then challenged with SLO for 2 h in the presence of PI. **h**, Percentage of PI+ neutrophils treated with IFN-β or IFN-γ (both 20 ng ml^{-1}) for 4 h and then challenged with PFO for 60 min in the presence of PI ($n = 3$). **i**, Flow cytometry plots of *S. aureus* phagocytosed by control or IFN-stimulated BMDMs. Macrophage cultures were stimulated with IFNs (100 ng ml^{-1}) for 24 h. BMDMs were then washed and incubated with PFO for 15 min. PFO-containing medium was then replaced with fresh medium containing pHrodo (red)-labeled *S. aureus*. Median fluorescence intensity (MFI) is indicated on the right. Data are representative of three independent experiments. Data in **a**–**h** are shown as the mean ± s.e.m. ($n = 3$). Statistical significance was determined using one-way ANOVA with Dunnett's correction (**a**–**e**), two-way ANOVA with Tukey's correction (**f** and **g**) or one-way repeated-measures ANOVA with Dunnett's correction (**h**). ***$P < 0.001$.

to IFN stimulation (Extended Data Fig. 3c). We also investigated whether IFN stimulation disrupted sphingomyelin-associated cholesterol in the plasma membrane. Ostreolysin A (OlyA) is a protein that specifically binds to cholesterol when associated with sphingomyelin[38]. To assess whether IFNs altered this pool of cholesterol in the plasma membrane of macrophages, we applied nanoscale secondary ion mass spectrometry (NanoSIMS)[39,40] using stable-isotope-labeled ALO-D4 and OlyA proteins. Consistent with the fluorescent ALO-D4 imaging studies, we observed a decrease in ^{15}N-labeled ALO-D4 binding in response to IFN-β stimulation (Fig. 3c). In contrast, [^{15}N]OlyA binding remained unchanged (Fig. 3c). High-throughput imaging studies with fluorescence-labeled OlyA showed a similar binding pattern for unstimulated and IFN-treated macrophage cultures (Extended Data Fig. 3d,e). Cholera toxin B (CTB) staining also showed no change, indicating that IFN signaling does not disrupt GM1 ganglioside-containing microdomains of the plasma membrane (Extended Data Fig. 3f). Together, these data support the conclusion that substantial amounts of cholesterol remain in the plasma membrane in association with sphingolipids after IFN stimulation[38]. Consistent with this concept, brief treatment of IFN-stimulated macrophages with sphingomyelinase[39] restored ALO-D4 binding (Fig. 3d,e). Thus, we conclude that CDCs are targeting a small fraction of non-sphingomyelin-associated cholesterol within the plasma membrane and that IFN signaling alters the availability of this pool of cholesterol to reduce CDC binding to phagocytes.

Cholesterol synthesis is linked to CDC susceptibility. We and others have previously reported that IFN signaling can reduce de novo cholesterol synthesis in macrophages[7,9,10]. Thus, we considered the possibility that changes in CDC sensitivity were linked to the cholesterol biosynthetic capacity of activated macrophages. To address this, we performed studies of [^{13}C] glucose tracer enrichment on PRR- and IFN-stimulated macrophages. IFN-β and IFN-γ stimulation markedly decreased macrophage cholesterol synthesis (Fig. 4a). Likewise, activation of STING or TLR3 also decreased cholesterol biosynthesis (Extended Data Fig. 4a,b). In contrast, activation of TLR1/TLR2, TLR7 and TLR9 increased the amounts of synthesized cholesterol accumulated (Extended Data Fig. 4b). These data largely mirror the differential sensitivity to CDCs that we observed in PRR- and IFN-stimulated macrophages, leading us to hypothesize that cholesterol biosynthetic flux determines susceptibility to CDC-induced plasma membrane damage. Consistent with this, pharmacologic inhibition of cholesterol synthesis using simvastatin reduced ALO-D4 binding to otherwise unstimulated macrophages and protected against CDC-mediated membrane damage (Fig. 4b,c

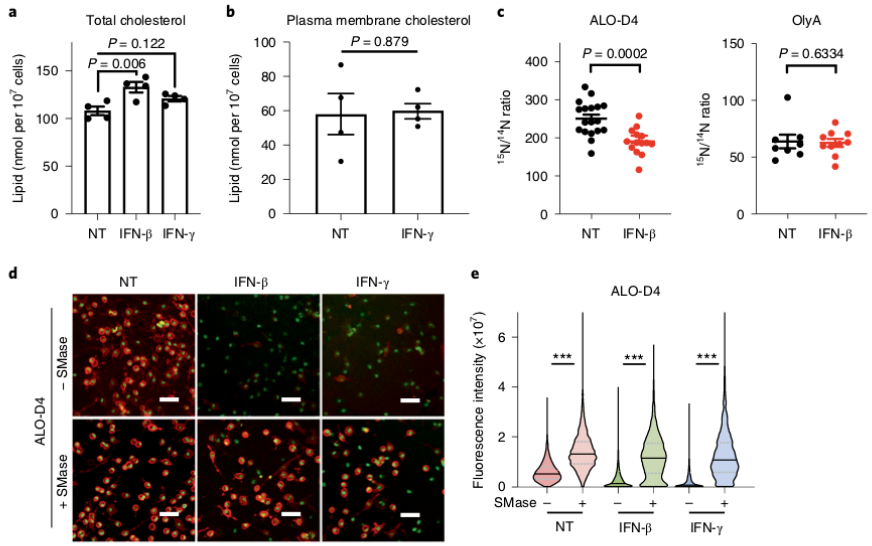


Fig. 3 | IFN signals reprogram cholesterol metabolism to decrease the pool of cholesterol targeted by CDCs. a, Total cholesterol (nmol lipid per 10^7 cells) from C57BL/6 BMDMs stimulated with IFN-β (20 ng ml^{-1}), IFN-γ (20 ng ml^{-1}) or unstimulated (NT) for 48 h ($n = 4$). **b**, Total plasma membrane cholesterol (nmol lipid per 10^7 cells) from C57BL/6 BMDMs stimulated with IFN-γ (40 ng ml^{-1}) or unstimulated (NT) for 24 h ($n = 4$). **c**, Quantification of [^{15}N]ALO-D4 or [^{15}N]OlyA binding on untreated (NT) or IFN-β-stimulated (100 ng ml^{-1}, 16 h) BMDMs determined by NanoSIMS. Quantification is based on the average ^{15}N/^{14}N ratio in cells. Data are the mean ± s.e.m ($n = 18$ and 14 for ALO-D4; $n = 8$ and 10 for OlyA). **d**, Confocal images of WT BMDM cultures stimulated with IFNs (20 ng ml^{-1}) for 24 h, treated with sphingomyelinase (SMase; 200 mU ml^{-1}) for 30 min at 37 °C and then stained with fluorescent ALO-D4 and DAPI. Scale bars represent 50 µm. **e**, Violin plots of cellular fluorescence intensity quantified from **d** ($n = 6{,}868$, 6,811, 3,404, 3,838, 2,930 and 2,642). Data are representative of three independent experiments (**a**, **d** and **e**), three independent samples (**c**) or from four biological replicates (**b**). Data in **a**–**c** are shown as the mean ± s.e.m. The violin plot in **e** is shown with the median (solid lines) and 25th and 75th percentiles (dashed lines). Statistical significance was determined using one-way ANOVA with Dunn's correction (**a**), an unpaired two-tailed Student's *t*-test (**b**) or a two-tailed Mann–Whitney test (**c** and **e**). ***$P < 0.001$.

markedly reduced ALO-D4 binding (Fig. 4f,g). Together, these data support the concept that the pool of newly synthesized cholesterol, rather than the total cholesterol available in the plasma membrane, is linked to the sensitivity or resistance of macrophages to CDC-mediated toxicity.

Production of 25HC is required for resistance to CDCs. IFNs have a well-defined role in inducing the conversion of cholesterol into 25HC via the enzyme CH25H in macrophages[5,13,14]. Thus, we next asked whether the endogenous production of 25HC is required for IFN-mediated protection against CDCs. To address this, control and CH25H-deficient macrophages were stimulated with IFNs and ALO-D4 binding was assessed. We observed that neither IFN-β nor IFN-γ treatment of CH25H-deficient macrophage cultures reduced ALO-D4 binding (Fig. 5a,b). Correspondingly, we found that stimulating CH25H-deficient macrophages with IFN-β or IFN-γ was unable to induce protection against either PFO or SLO challenge (Fig. 5c,d and Extended Data Fig. 5a,b). We also observed that, in contrast to control macrophages, IFN treatment of CH25H-deficient macrophages reduced their ability to maintain efficient phagocytosis of bacteria when challenged with PFO (Fig. 5e). Importantly, provisioning of 25HC to cultures of CH25H-deficient or control macrophages protected them from CDC-induced pore formation (Fig. 5f) and largely restored their ability to phagocytose bacteria (Fig. 5e). Thus, endogenous production of 25HC appears to be critical for macrophage protection to CDCs.

Next, we sought to better define the mechanism by which IFN signaling and the production of 25HC mediate protection against CDCs. Our studies to this point correlated a decrease in cholesterol biosynthesis with protection against challenge by CDCs, leading us to ask whether the loss of CH25H altered IFN-mediated reduction in the cholesterol synthesis. [^{13}C] glucose tracer enrichment studies showed that loss of CH25H in macrophages resulted in basally higher production of cholesterol as compared to control macrophages but did not alter total cholesterol content (Fig. 5g). IFN-β stimulation was able to reduce cholesterol synthesis in both control and CH25H-deficient macrophages, but the reduction was attenuated in the absence of CH25H, resulting in a synthetic profile that more closely resembled that observed in unstimulated control macrophages (Fig. 5g). Treatment of CH25H-deficient macrophages with simvastatin was sufficient to reduce ALO-D4 binding, supporting the concept that the ability of 25HC to regulate cholesterol synthetic flux is an important mechanism by which IFNs mediate protection from CDCs (Fig. 5h,i).

Cholesterol esterification contributes to CDC resistance of macrophages. We observed that IFNs induce a number of genes involved in cholesterol esterification and efflux (Fig. 6a), prompting us to ask whether these processes were also necessary for IFN-mediated protection from CDCs. As expected, IFN treatment of macrophages increased cholesterol ester formation[43,44] (Fig. 6b). Unexpectedly, we found that cholesterol in culture medium was not required for accumulation of esterified cholesterol in response to IFNs (Extended Data Fig. 6a), indicating that the source of cholesterol for esterification can be derived from host cell membranes rather than from the extracellular milieu. Pharmacologic inhibition of the cholesterol

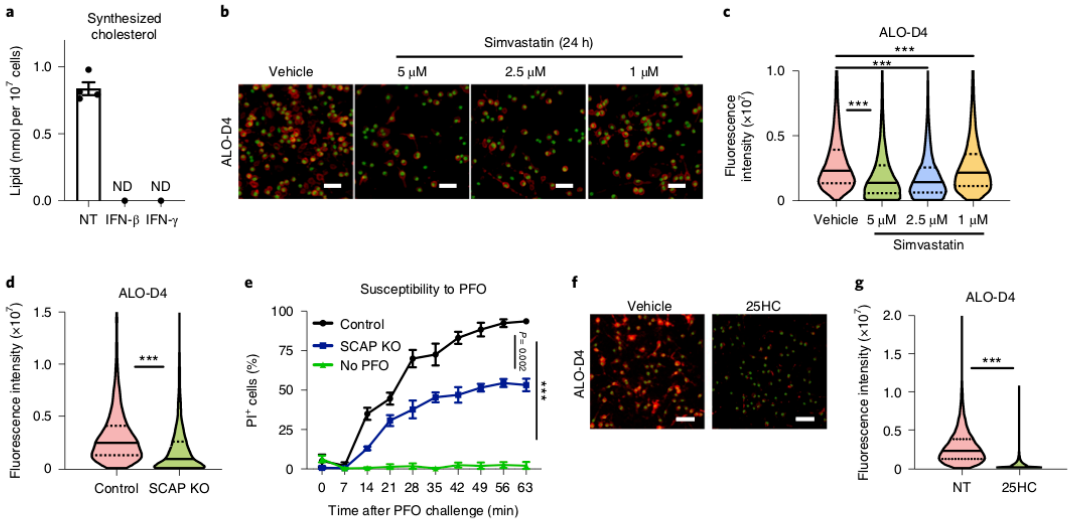


Fig. 4 | Cholesterol synthesis is linked to CDC susceptibility. a, Net synthesized cholesterol (nmol lipid per 10^7 cells) from C57BL/6 BMDMs stimulated with IFN-β (20 ng ml^{-1}), IFN-γ (20 ng ml^{-1}) or unstimulated (NT) for 48 h. Synthesized cholesterol was determined by GC–MS and ISA modeling ($n = 4$). ND, not detectable. **b**, Confocal images of WT BMDMs treated with simvastatin (1–5 μM) for 24 h and then stained with fluorescent ALO-D4 and DAPI. **c**, Violin plots of cellular fluorescence intensity quantified from **b** ($n = 5{,}713$, 2,252, 2,808 and 5,556). **d**, Violin plots of cellular fluorescence intensity quantified from control or SCAP KO BMDMs stained with fluorescent ALO-D4 and DAPI ($n = 3{,}306$ and 2,927). **e**, Percentage of PI$^+$ control or SCAP KO BMDMs challenged with PFO for 60 min in the presence of PI ($n = 3$). **f**, Confocal images of WT BMDM cultures incubated with 25HC (3 μM) for 4 h and then stained with fluorescent ALO-D4 and DAPI. **g**, Violin plots of cellular fluorescence intensity quantified from **f** ($n = 5{,}861$ and 5,769). Data are representative of three independent experiments. Data in **a** and **e** are shown as the mean ± s.e.m. Violin plots in **c**, **d** and **g** are shown with the median (solid lines) and 25th and 75th percentiles (dashed lines). Statistical significance was determined using a Kruskal–Wallis test with Dunn's correction (**c**), a two-tailed Mann–Whitney test (**d** and **g**) or one-way repeated-measures ANOVA (**e**). ***$P < 0.001$. Scale bars in **b** and **f** represent 50 μm.

esterification enzymes acyl-CoA:cholesterol acyltransferase 1 and 2 (ACAT1 and ACAT2) partially attenuated the protection from PFO-mediated pore formation conveyed by IFN signaling (Fig. 6c and Extended Data Fig. 6b). 25HC has also been shown to mediate cholesterol ester formation[45,46]; thus, we asked whether CH25H deficiency influenced cholesterol ester production in response to IFN signals. We found that quiescent CH25H-deficient macrophages had modestly increased levels of cholesterol esters and were able to increase this pool to the levels seen in control macrophages in response to IFN stimulation (Extended Data Fig. 6c). Consistent with this, we observed that treatment of CH25H-deficient macrophages with an ACAT inhibitor further increased ALO-D4 binding and increased their susceptibility to CDC-mediated membrane damage (Fig. 6d and Extended Data Fig. 6d). These data indicate that synthesis of cholesterol esters in macrophages is required, in part, for resistance to CDCs.

To test whether cholesterol efflux was required for IFN-mediated protection from CDCs, we activated macrophages deficient in the key cholesterol efflux transporters ATP-binding cassette sub-family A member 1 (ABCA1) or ATP-binding cassette sub-family G member 1 (ABCG1) with IFN[11,12]. We found that IFN treatment (24 h) reduced ALO-D4 binding in both ABCA1- and ABCG1-deficient macrophages to a level similar to that of control macrophages (Extended Data Fig. 6e). No difference in susceptibility to CDC-mediated pore formation was observed between quiescent control and ABCA1-deficient macrophages (Fig. 6e). Furthermore, IFN treatment conveyed resistance to PFO challenge for ABCA1-deficient macrophages, indicating that efflux was unlikely to be an important component of the IFN-mediated protection under these conditions. However, we did find that pharmacological enforcement of cholesterol efflux by treating with liver X receptors (LXR) agonist GW3965 (ref. [47]) for 24 h decreased ALO-D4 binding and provided robust protection against CDCs in wild-type macrophages (Fig. 6f,g). We also observed that LXR activation with GW3965 decreased ALO-D4 binding and conveyed protection for CH25H-deficient macrophages when challenged with PFO (Fig. 6f,g), indicating that 25HC mediates protection by regulating synthesis of cholesterol rather than by blocking CDCs at the membrane or inducing esterification. Taken together, these data support a model where IFNs reduce cholesterol synthesis and increase esterification of free cholesterol to reduce the cholesterol pool size in the plasma membrane targeted by CDCs.

25HC mediates protection to CDC-induced tissue damage. Finally, we asked whether reprogramming of cholesterol homeostasis would influence CDC susceptibility in vivo. To this end, SLO was injected into the skin of control and CH25H-deficient mice. Injection sites were visually monitored for up to 48 h for the appearance of erythema and ulceration. The SLO challenge resulted in modest erythema within 24 h and small ulcers at 48 h (Fig. 7a and Extended Data Fig. 7a). In contrast, CH25H-deficient mice developed severe erythema within 24 h, which progressed to ulcerative lesions that had upwards of a six-fold larger area than those of their wild-type counterparts (Fig. 7a and Extended Data Fig. 7a). Histologic examination confirmed greater epidermal ulceration area, along with a greater depth of tissue necrosis into the deep dermis and, in some cases, the adipose tissue, in CH25H-deficient mice (Fig. 7b). We also asked whether injecting 25HC into the dermis would protect mice from CDC-induced tissue damage. For these studies, we used the ALO challenge,

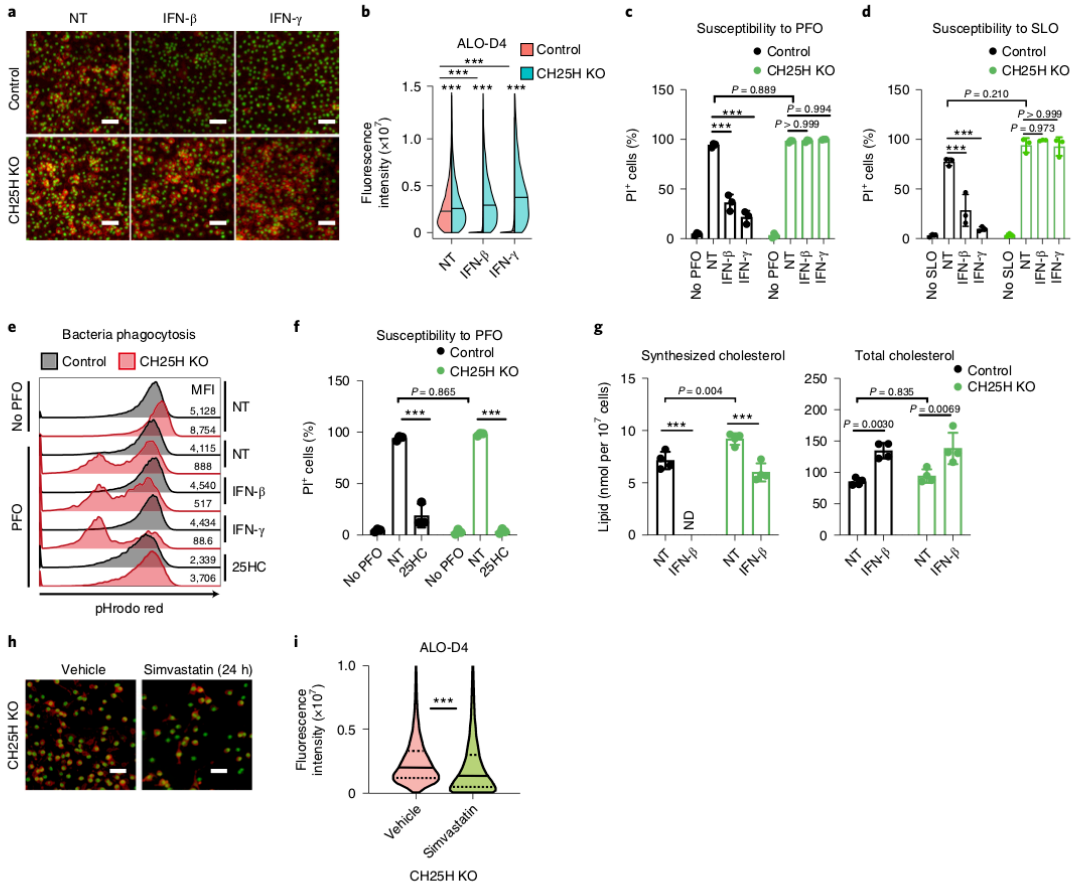


Fig. 5 | Production of 25HC is required to maintain changes in plasma membrane cholesterol and mediates resistance to CDCs. a, Confocal images of control or CH25H-deficient (CH25H KO) BMDM cultures stimulated with IFN-β or IFN-γ (20 ng ml^{-1}) for 24 h and then stained with fluorescent ALO-D4 and DAPI. **b**, Violin plots of cellular fluorescence intensity quantified from **a** (n = 5,294, 5,505, 4,564, 4,066, 5,370 and 4,962). **c**, Percentage of PI$^+$ control or CH25H KO BMDMs stimulated with IFNs (20 ng ml^{-1}) for 24 h and then challenged with PFO for 60 min in the presence of PI. **d**, Percentage of PI$^+$ control or CH25H KO BMDMs stimulated with IFNs (20 ng ml^{-1}) for 24 h and then challenged with SLO for 2 h in the presence of PI. **e**, Flow cytometry plots of *S. aureus* phagocytosed by control or CH25H KO BMDMs. Macrophage cultures were treated with IFNs (100 ng ml^{-1}) or 25HC (3 μM). After 24 h, BMDMs were washed and then incubated with PFO for 15 min. PFO-containing medium was then replaced with fresh medium containing pHrodo (red)-labeled *S. aureus*. MFI is indicated on the right. **f**, Percentage of PI$^+$ control or CH25H KO BMDMs incubated with 25HC (1 μM) overnight and then challenged with PFO for 60 min in the presence of PI. **g**, Net synthesized and total cholesterol (nmol per 10^7 cells) from CH25H KO or control BMDMs stimulated with IFN-β (20 ng ml^{-1}) or unstimulated (NT) for 48 h. Total and synthesized cholesterol was determined by GC–MS and ISA modeling. ND, not detectable. **h**, Confocal images of CH25H KO BMDMs treated with simvastatin (5 μM) for 24 h and then stained with fluorescent ALO-D4 and DAPI. **i**, Violin plots of cellular fluorescence intensity quantified from **h** (n = 6,666 and 2,867). Data are representative of three independent experiments. Data in **c**, **d**, **f** and **g** are shown as the mean ± s.e.m. (n = 3 in **c**, **d** and **f**; n = 4 in **g**). Violin plots in **b** and **i** are shown with the median (solid lines) and 25th and 75th percentiles (dashed lines in **i**). Statistical significance was determined using a Kruskal–Wallis test with Dunn's correction (**b**), two-way ANOVA with Tukey's correction (**c**, **d**, **f** and **g**) or a two-tailed Mann–Whitney test (**i**). ***P < 0.001. Scale bars in **a** and **h** represent 50 μm.

because we found that ALO reliably induces larger skin lesions in control mice compared to SLO. We found that preinjection of 25HC into the dermis of mice for 6 h markedly reduced ALO-induced tissue damage (Fig. 7c and Extended Data Fig. 7b). Histologic examination revealed that, in general, 25HC pretreatment completely prevented tissue damage, or, in cases with damage, 25HC pretreatment reduced both ulcer area and depth (Fig. 7d). Thus, we conclude that generation of 25HC and the consequent ability to reprogram cholesterol metabolism protect mice from CDC-mediated tissue damage.

Discussion

Immune cells have presumably evolved strategies to evade the deleterious effects of virulence factors to efficiently combat pathogens and protect the host from damage. In this study, we delineate a mechanism by which IFN signals alter the abundance of a small

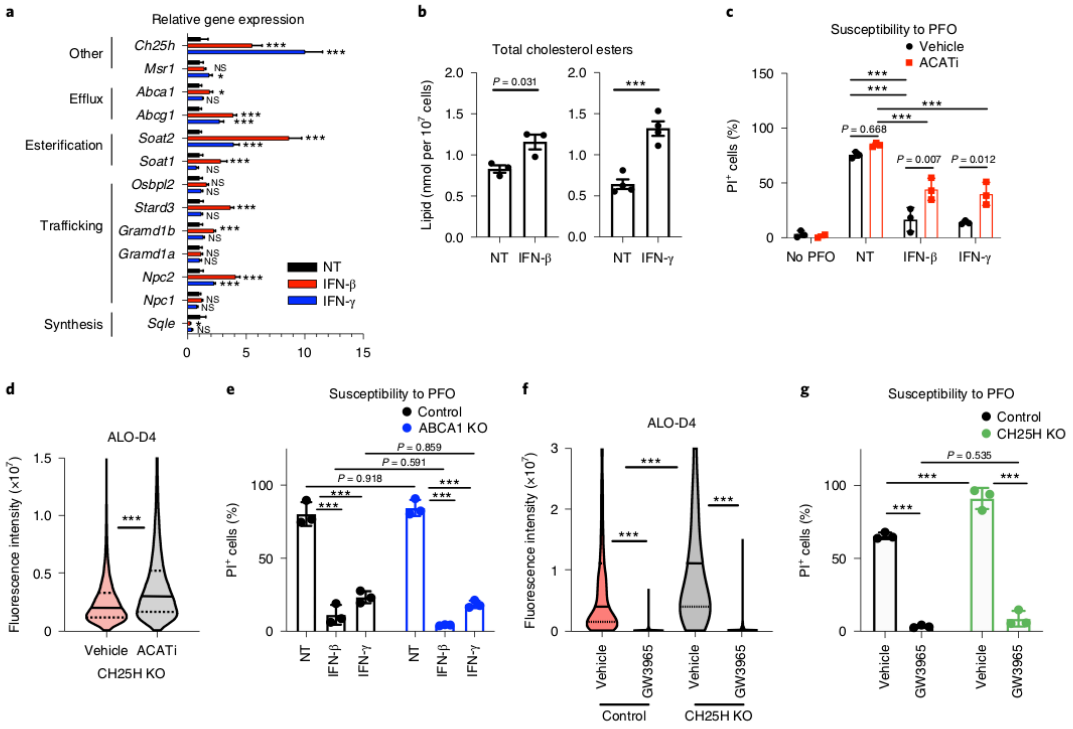


Fig. 6 | Cholesterol esterification contributes to CDC resistance of macrophages. a, Quantitative PCR (qPCR) analysis of genes that regulate cholesterol homeostasis in WT BMDMs treated with IFNs (20 ng ml^{-1}) for 24 h. **b**, Quantification (nmol lipid per 10^7 cells) of total cholesterol ester in WT BMDMs stimulated with IFN-β (20 ng ml^{-1}) or IFN-γ (20 ng ml^{-1}) for 48 h. Cholesterol ester pool sizes were determined by direct infusion mass spectrometry (left, $n = 3$; right, $n = 4$). **c**, Percentage of PI$^+$ BMDMs treated with IFN-β or IFN-γ (20 ng ml^{-1}) for 24 h in the presence of the ACAT inhibitor (ACATi) 58-035 (4.3 μM) and then challenged with PFO for 60 min in the presence of PI. **d**, Violin plots of cellular fluorescence intensity quantified from control or CH25H KO BMDMs treated with ACAT inhibitor 58-035 (4.3 μM) for 24 h and then stained with fluorescent ALO-D4 and DAPI ($n = 6{,}666$ and 9,997). **e**, Percentage of PI$^+$ control or ABCA1 KO BMDMs stimulated with IFNs (20 ng ml^{-1}) for 24 h and then challenged with PFO for 60 min in the presence of PI. **f**, Violin plots of cellular fluorescence intensity quantified from control or CH25H KO BMDMs treated with LXR agonist GW3965 (1 μM) for 24 h and then stained with fluorescent ALO-D4 and DAPI ($n = 2{,}304$, 3,954, 2,712 and 2,322). **g**, Percentage of PI$^+$ control or CH25H KO BMDMs treated with LXR agonist GW3965 (1 μM) for 24 h and then challenged with PFO for 60 min in the presence of PI. Data are representative of three independent experiments. Data in **a**–**c**, **e** and **g** are shown as the mean ± s.e.m. ($n = 3$) unless otherwise specified. Violin plots in **d** and **f** are shown with the median (solid lines) and 25th and 75th percentiles (dashed lines). Statistical significance was determined using one-way ANOVA with Dunnett's correction (**a**), an unpaired two-tailed Student's *t*-test (**b**), two-way ANOVA with Tukey's correction (**c**, **e** and **g**), a two-tailed Mann–Whitney test (**d**) or a Kruskal–Wallis test with Dunn's correction (**f**). ***$P < 0.001$; NS, not significant.

pool of cholesterol in the plasma membrane to protect macrophages and neutrophils from pathogen-produced toxins. It is likely that IFN-mediated reprogramming of cholesterol homeostasis will extend into other cell types, such as endothelium, adipocytes and epithelium, given the extent of tissue damage observed in the skin of CH25H-deficient mice in response to CDC challenge. Mechanistically, we find that inhibition of new cholesterol synthesis through the actions of CH25H is required to induce this protection mechanism. We also find that the IFN-mediated conversion of cholesterol into cholesterol esters is required for maximal protection. In combination, IFN regulation of these distinct but interrelated cholesterol metabolic pathways ensures that the cholesterol pool targeted by CDCs remains small, and consequently macrophages and neutrophils retain their function. These data also imply that pharmacologic manipulation of the cholesterol homeostatic machinery in the skin might decrease tissue damage associated with the deleterious effects of Gram-positive infections observed in necrotizing fasciitis or similar soft tissue infections. However, it will be necessary to determine whether manipulating cholesterol homeostasis in infected tissues interferes with other host defense pathways and clearance of microorganisms.

One unexpected finding of our studies is that the IFN-induced changes in plasma membrane cholesterol are restricted to a very small fraction of the total membrane cholesterol. The exact pool of cholesterol in the plasma membrane targeted by CDCs remains poorly characterized[19,20], and our data suggest that a minor fraction of cholesterol in the plasma membrane is required for CDC binding and pore formation. The majority of cholesterol is not evenly distributed across the plasma membrane[2,48], but rather is tightly associated with sphingomyelins or other phospholipids in microdomains[38,49]. Recent studies provide evidence for distinct pools of cholesterol within the plasma membrane. These studies describe a

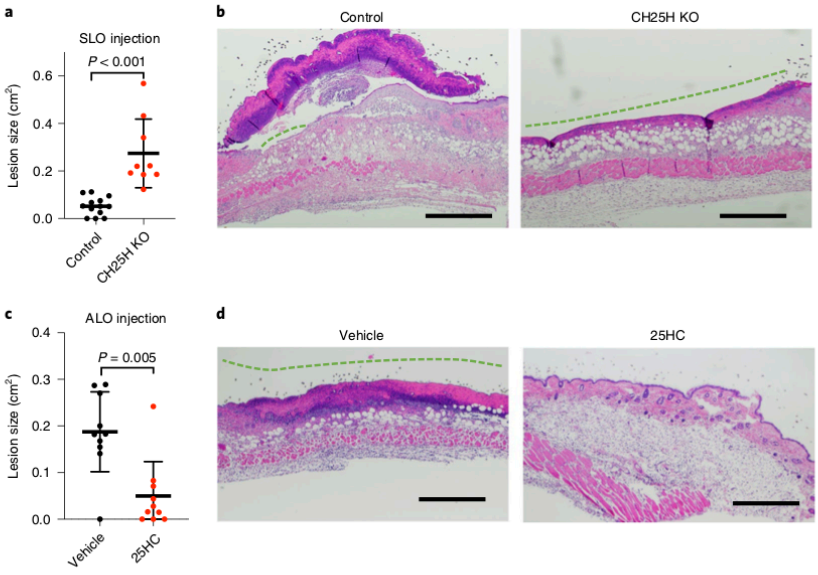


Fig. 7 | 25HC mediates protection against CDC-induced tissue damage. a, Quantification of lesion size from WT ($n=13$) and CH25H KO ($n=9$) mice challenged with SLO for 48 h. **b**, Representative histology from mice demonstrates serum crust overlying partial thickness erosion with a small full-thickness ulceration (green dashed line) of the epidermis and a robust inflammatory response in the dermis and adipose tissue in control (C57BL/6) mice (left). Wound areas in CH25H-deficient mice demonstrate a large ulceration (green dashed line) with necrosis of the papillary and much of the reticular dermis, with a less pronounced inflammatory response (right). **c**, Quantification of lesion size from WT mice ($n=10$ each) pretreated with vehicle or 25HC for 6 h and challenged with ALO for 48 h. **d**, Representative histology of wounds demonstrates full-thickness ulceration (green dashed line) with overlying neutrophilic/serum crust by ALO in wounds pretreated with ethanol vehicle (left). Edema and a mixed mononuclear and neutrophilic inflammatory host response without ulceration are present in the dermis and adipose tissue of mice pretreated with 25HC (right). Data in **a** and **c** are shown as the mean ± s.e.m. Statistical significance was determined using a two-tailed Mann–Whitney test. Scale bars represent 500 μm (**b** and **d**).

small pool of cholesterol, termed 'accessible' cholesterol, that is in rapid equilibrium with the endoplasmic reticulum cholesterol pool and, consequently, the cholesterol biosynthetic machinery[48]. We suspect that this is the cholesterol pool that is targeted by CDCs and rapidly reduced upon IFN signaling. Our data indicate that this IFN-dependent reduction of the accessible cholesterol pool occurs through the combined actions of the ACAT enzymes and inhibition of cholesterol biosynthesis. We cannot rule out the possibility that IFN signals also move cholesterol laterally or vertically from the accessible pool into other domains in the plasma membrane. It will be of considerable interest to determine the spatial distribution of the accessible cholesterol pool within the plasma membrane and how this specific cholesterol pool relates to canonical domains in the plasma membrane as described previously.

We also believe that other cholesterol metabolic pathways will be involved in this protective mechanism. We found that IFNs induce changes in ALO-D4 binding within 2 h; thus, it seems likely that direct signaling events downstream of the IFN receptors initiate this process, and it will be of interest to define the cholesterol transport machinery involved in the rapid removal of this small cholesterol pool from the plasma membrane. The observation that both type I and type II IFNs regulate a broad cassette of genes involved in cellular cholesterol homeostasis (for example, efflux, esterification, synthesis and intracellular trafficking) leads us to surmise that considerable pressure exists to sequester or limit accessible or free cholesterol from pathogens, and it will be of interest to determine the pathophysiologic context where this reshaping of membrane cholesterol homeostasis contributes to host defense. These observations also imply that other mechanisms that can influence the pool size of accessible cholesterol, such as metabolic dysfunction, could influence the susceptibility of cells to these pore-forming toxins. Consistent with this notion, it has been shown that diabetes and obesity are independent risk factors for necrotizing fasciitis[50,51], and we posit that restoring cellular cholesterol homeostasis with drugs that target cholesterol metabolic machinery might serve to decrease the severity of disease in these patients. In support of this idea, we find that enforced cholesterol efflux through the activation of LXR can induce protection independently of IFN signaling. Likewise, injecting 25HC into the skin protected the dermis from ALO-induced damage, providing initial proof of concept that metabolic manipulation could be an adjunctive approach to conventional antimicrobials.

In conclusion, these studies provide evidence that rapid remodeling of the lipid content of membranes is used as a host defense strategy by macrophages and neutrophils. Given the breadth of microbial and viral proteins that target host lipids, we propose that signal-specific reprogramming of the lipidome of cells will be a fundamentally important strategy for host cells to evade different pathogens.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of

data and code availability are available at https://doi.org/10.1038/s41590-020-0695-4.

Received: 11 March 2020; Accepted: 28 April 2020;
Published online: 8 June 2020

References

1. Lange, Y., Swaisgood, M. H., Ramos, B. V. & Steck, T. L. Plasma membranes contain half the phospholipid and 90% of the cholesterol and sphingomyelin in cultured human fibroblasts. *J. Biol. Chem.* **264**, 3786–3793 (1989).
2. Ikonen, E. Cellular cholesterol trafficking and compartmentalization. *Nat. Rev. Mol. Cell Biol.* **9**, 125–138 (2008).
3. van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9**, 112–124 (2008).
4. Luo, J., Yang, H. & Song, B. L. Mechanisms and regulation of cholesterol homeostasis. *Nat. Rev. Mol. Cell Biol.* **21**, 225–245 (2020).
5. Blanc, M. et al. The transcription factor STAT-1 couples macrophage synthesis of 25-hydroxycholesterol to the interferon antiviral response. *Immunity* **38**, 106–118 (2013).
6. Reboldi, A. et al. 25-hydroxycholesterol suppresses interleukin-1-driven inflammation downstream of type I interferon. *Science* **345**, 679–684 (2014).
7. Dang, E. V., McDonald, J. G., Russell, D. W. & Cyster, J. G. Oxysterol restraint of cholesterol synthesis prevents AIM2 inflammasome activation. *Cell* **171**, 1057–1071 (2017).
8. Araldi, E. et al. Lanosterol modulates TLR4-mediated innate immune responses in macrophages. *Cell Rep.* **19**, 2743–2755 (2017).
9. Blanc, M. et al. Host defense against viral infection involves interferon-mediated down-regulation of sterol biosynthesis. *PLoS Biol.* **9**, e1000598 (2011).
10. York, A. G. et al. Limiting cholesterol biosynthetic flux spontaneously engages type I IFN signaling. *Cell* **163**, 1716–1729 (2015).
11. Shibata, N. & Glass, C. K. Regulation of macrophage function in inflammation and atherosclerosis. *J. Lipid Res.* **50** (Suppl.), S277–S281 (2009).
12. Tall, A. R. & Yvan-Charvet, L. Cholesterol, inflammation and innate immunity. *Nat. Rev. Immunol.* **15**, 104–116 (2015).
13. Bauman, D. R. et al. 25-hydroxycholesterol secreted by macrophages in response to Toll-like receptor activation suppresses immunoglobulin A production. *Proc. Natl Acad. Sci. USA* **106**, 16764–16769 (2009).
14. Liu, S.-Y. et al. Interferon-inducible cholesterol-25-hydroxylase broadly inhibits viral entry by production of 25-hydroxycholesterol. *Immunity* **38**, 92–105 (2013).
15. Viard, M. et al. Role of cholesterol in human immunodeficiency virus type 1 envelope protein-mediated fusion with host cells. *J. Virol.* **76**, 11584–11595 (2002).
16. Goluszko, P. & Nowicki, B. Membrane cholesterol: a crucial molecule affecting interactions of microbial pathogens with mammalian cells. *Infect. Immun.* **73**, 7791–7796 (2005).
17. Rawat, S. S. et al. Modulation of entry of enveloped viruses by cholesterol and sphingolipids. *Mol. Membr. Biol.* **20**, 243–254 (2003).
18. Mazzon, M. & Mercer, J. Lipid interactions during virus entry and infection. *Cell. Microbiol.* **16**, 1493–1502 (2014).
19. Tweten, R. K. Cholesterol-dependent cytolysins, a family of versatile pore-forming toxins. *Infect. Immun.* **73**, 6199–6209 (2005).
20. Tweten, R. K., Hotze, E. M. & Wade, K. R. The unique molecular choreography of giant pore formation by the cholesterol-dependent cytolysins of Gram-positive bacteria. *Annu. Rev. Microbiol.* **69**, 323–340 (2015).
21. Gilbert, R. J. Inactivation and activity of cholesterol-dependent cytolysins: what structural studies tell us. *Structure* **13**, 1097–1106 (2005).
22. Timmer, A. M. et al. Streptolysin O promotes group A *Streptococcus* immune evasion by accelerated macrophage apoptosis. *J. Biol. Chem.* **284**, 862–871 (2009).
23. Bhattacharjee, P. & Keyel, P. A. Cholesterol-dependent cytolysins impair pro-inflammatory macrophage responses. *Sci. Rep.* **8**, 6458 (2018).
24. Corrotte, M., Fernandes, M. C., Tam, C. & Andrews, N. W. Toxin pores endocytosed during plasma membrane repair traffic into the lumen of MVBs for degradation. *Traffic* **13**, 483–494 (2012).
25. McNeil, P. L. & Kirchhausen, T. An emergency response team for membrane repair. *Nat. Rev. Mol. Cell Biol.* **6**, 499–505 (2005).
26. Romero, M. et al. Intrinsic repair protects cells from pore-forming toxins by microvesicle shedding. *Cell Death Differ.* **24**, 798–808 (2017).
27. Oishi, Y. et al. SREBP1 contributes to resolution of pro-inflammatory TLR4 signaling by reprogramming fatty acid metabolism. *Cell Metab.* **25**, 412–427 (2017).
28. Stetson, D. B. & Medzhitov, R. Type I interferons in host defense. *Immunity* **25**, 373–381 (2006).
29. McNab, F., Mayer-Barber, K., Sher, A., Wack, A. & O'Garra, A. Type I interferons in infectious disease. *Nat. Rev. Immunol.* **15**, 87–103 (2015).
30. Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. *Cell* **140**, 805–820 (2010).
31. Woodward, J. J., Iavarone, A. T. & Portnoy, D. A. c-di-AMP secreted by intracellular *Listeria monocytogenes* activates a host type I interferon response. *Science* **328**, 1703–1705 (2010).
32. Pandey, A. K. et al. NOD2, RIP2 and IRF5 play a critical role in the type I interferon response to *Mycobacterium tuberculosis*. *PLoS Pathog.* **5**, e1000500 (2009).
33. Gay, A., Rye, D. & Radhakrishnan, A. Switch-like responses of two cholesterol sensors do not require protein oligomerization in membranes. *Biophys. J.* **108**, 1459–1469 (2015).
34. Chakrabarti, R. S. et al. Variability of cholesterol accessibility in human red blood cells measured using a bacterial cholesterol-binding toxin. *Elife* **6**, e23355 (2017).
35. Infante, R. E. & Radhakrishnan, A. Continuous transport of a small fraction of plasma membrane cholesterol to endoplasmic reticulum regulates total cellular cholesterol. *Elife* **6**, e25466 (2017).
36. Endapally, S., Infante, R. E. & Radhakrishnan, A. Monitoring and modulating intracellular cholesterol trafficking using ALOD4, a cholesterol-binding protein. *Methods Mol. Biol.* **1949**, 153–163 (2019).
37. Maxfield, F. R. & Wustner, D. Analysis of cholesterol trafficking with fluorescent probes. *Methods Cell Biol.* **108**, 367–393 (2012).
38. Endapally, S. et al. Molecular discrimination between two conformations of sphingomyelin in plasma membranes. *Cell* **176**, 1040–1053 (2019).
39. He, C. et al. High-resolution imaging and quantification of plasma membrane cholesterol by NanoSIMS. *Proc. Natl Acad. Sci. USA* **114**, 2000–2005 (2017).
40. He, C. et al. Macrophages release plasma membrane-derived particles rich in accessible cholesterol. *Proc. Natl Acad. Sci. USA* **115**, E8499–E8508 (2018).
41. Kandutsch, A. A. & Chen, H. W. Regulation of sterol synthesis in cultured cells by oxygenated derivatives of cholesterol. *J. Cell Physiol.* **85**, 415–424 (1975).
42. Goldstein, J. L., DeBose-Boyd, R. A. & Brown, M. S. Protein sensors for membrane sterols. *Cell* **124**, 35–46 (2006).
43. Panousis, C. G. & Zuckerman, S. H. Regulation of cholesterol distribution in macrophage-derived foam cells by interferon-γ. *J. Lipid Res.* **41**, 75–83 (2000).
44. Keyel, P. A., Tkacheva, O. A., Larregina, A. T. & Salter, R. D. Coordinate stimulation of macrophages by microparticles and TLR ligands induces foam cell formation. *J. Immunol.* **189**, 4621–4629 (2012).
45. Collins, J. L. et al. Identification of a nonsteroidal liver X receptor agonist through parallel array synthesis of tertiary amines. *J. Med. Chem.* **45**, 1963–1966 (2002).
46. Das, A., Brown, M. S., Anderson, D. D., Goldstein, J. L. & Radhakrishnan, A. Three pools of plasma membrane cholesterol and their relation to cholesterol homeostasis. *Elife* **3**, e02882 (2014).
47. Pike, L. J. Lipid rafts: bringing order to chaos. *J. Lipid Res.* **44**, 655–667 (2003).
48. Das, D. K., Baker, M. G. & Venugopal, K. Risk factors, microbiological findings and outcomes of necrotizing fasciitis in New Zealand: a retrospective chart review. *BMC Infect. Dis.* **12**, 348 (2012).
49. Arif, N., Yousfi, S. & Vinnard, C. Deaths from necrotizing fasciitis in the United States, 2003–2013. *Epidemiol. Infect.* **144**, 1338–1344 (2016).
50. Timmins, J. M. et al. Targeted inactivation of hepatic *Abca1* causes profound hypoalphalipoproteinemia and kidney hypercatabolism of apoA-I. *J. Clin. Invest.* **115**, 1333–1342 (2005).
51. Sag, D., Cekic, C., Wu, R., Linden, J. & Hedrick, C. C. The cholesterol transporter ABCG1 links cholesterol homeostasis and tumour immunity. *Nat. Commun.* **6**, 6354 (2015).

Methods

Mouse strains. Wild-type and gene-targeted mice were purchased from The Jackson Laboratory: WT C57BL/6J (JAX 000664), B6.129S6-*Ch25h*^{tm1Rus}/J (*Ch25h*$^{-/-}$, JAX 016263) and B6(Cg)-*Ifnar1*$^{tm1.2Es}$/J (*Ifnar1*$^{-/-}$, JAX 028288). SCAP-deficient macrophages were generated from a *Lzy2*-Cre×*Scap*$^{fl/fl}$ cross as previously described[10]. ABCA1- and ABCG1-deficient macrophages were generated from *Lzy2*-Cre×*Abca1*$^{fl/fl}$ or *Lyz2*-Cre×*Abcg1*$^{fl/fl}$ animals as previously described[50,51]. All mice used for BMDMs were males aged 8–16 weeks. For in vivo CDC injections, mice were female and aged 8–12 weeks. All mice were maintained in pathogen-free facilities of the University of California, Los Angeles. All experiments on mice and tissues collected from mice were performed in strict accordance with the University of California policy on the humane and ethical treatment of animals.

PBMC-derived macrophages. Human monocyte-derived monocytes were isolated from leukopacks using standard Ficoll isolation procedures and plastic adherence. Isolation of monocytes was conducted by the UCLA Center for AIDS Research Virology Core laboratory. Monocytes were differentiated into macrophages with 50 ng ml^{-1} human GM-CSF (Peprotech, 300-03) in IMDM medium (HyClone) with 10% FBS and 1% (vol/vol) penicillin–streptomycin for 7 d before experimental use.

Mouse cells. Bone marrow cells were differentiated into macrophages in DMEM containing 10% (vol/vol) FBS (HyClone, GE SH3007103), 5% (vol/vol) M-CSF conditioned medium, 1% (vol/vol) penicillin–streptomycin, 1% (vol/vol) glutamine (Invitrogen) and 0.5% (vol/vol) sodium pyruvate (Invitrogen) for 7–9 d before experimental use. Cells were changed to medium with 5% FBS at the time of stimulation. Mouse neutrophils were isolated from the bone marrow using the EasySep Mouse Neutrophil Enrichment Kit following the manufacturer's protocol (StemCell, 19762).

Reagents. PRR ligands used were LPS (Invivogen, tlrl-smlps), poly(I:C) (Invivogen, tlrl-pic (HMW)), Pam3CSK4 (Invivogen, tlrl-pms), CL307 (Invivogen, tlrl-c307), ODN1668 (Invivogen, tlrl-1668), 2′,3′-cGAMP (Invivogen, tlrl-nacga23), c-di-GMP (Invivogen, tlrl-cdg), 5′-ppp-dsRNA (Invivogen, tlrl-3prn), iE-DAP (Invivogen, tlrl-dap) and *N*-glycolyl-MDP (Invivogen, tlrl-gmdp). Cytokines used were recombinant murine IFN-γ (Peprotech, 315-05), recombinant mouse IFN-β1 (carrier-free; BioLegend, 581302), recombinant human IFN-β (Peprotech, 300-02BC) and IFN-α was a kind gift from the laboratory of R. Modlin (UCLA). HyClone IMDM medium (16750-088) and FBS (SH3007103) were purchased from VWR. PFO was a kind gift from the laboratory of A.S. Divakaruni[52] (Agilent, 102504-100). SLO from *Streptococcus pyogenes* (gamma irradiated; S0149-25KU), 58-035 (S9318), simvastatin (S6196-5MG) and LXR agonist GW3965 (G6295-5MG) were purchased from Sigma-Aldrich.

Cholesterol ester analysis. Macrophages were cultured in six-well dishes (Fisher, 08-772-1B) and stimulated with TLR ligands as described above. At 48 h after stimulation, cells were imaged for cell counting as previously described[10], scraped and spun down in PBS and snap frozen as cell pellets. A modified Bligh and Dyer extraction[53] was carried out on samples. Before biphasic extraction, a 13-lipid class Lipidyzer Internal Standard Mix was added to each sample (AB SCIEX, 5040156). Following two successive extractions, pooled organic layers were dried down in a Genevac EZ-2 Elite. Lipid samples were resuspended in 1:1 methanol/dichloromethane with 10 mM ammonium acetate and transferred to robovials (Thermo Fisher, 10800107) for analysis. Samples were analyzed on the Lipidyzer Platform for targeted quantitative measurement of 1,100 lipid species across 13 classes. The Differential Mobility device on the Lipidyzer was tuned with a SelexION tuning kit (SCIEX, 5040141). Instrument settings, tuning settings and the MRM list are available upon request. Data analysis was performed on Lipidyzer software. Quantitative values were normalized to cell counts.

Isotope-enrichment experiments. BMDMs differentiated for 8 d were transferred to complete medium containing 50% [^{13}C] glucose (Cambridge Isotope Laboratories, CLM-1396-MPT-PK) with or without TLR stimulation for 48 h before collection. Analysis of labeled fatty acids and cholesterol was performed as described previously[10,54,55]. The relative contributions of synthesis to the total cholesterol pool over the 48-h labeling period were determined by fitting the isotopolog distributions for cholesterol in a model similar to isotopomer spectral analysis (ISA) as described previously[10,54,55].

Gene expression analysis. RNA was extracted from all cells with TRIzol (Thermo Fisher, 15596-018) using the manufacturer's protocols. cDNA was synthesized with a high-capacity cDNA reverse transcription kit (Applied Biosystems, 4368814) according to the manufacturer's instructions (700 ng µl^{-1} RNA per cDNA synthesis reaction). qPCR was conducted on the Roche LightCycler 480 using SYBR green master mix (Kapa Biosciences) or PowerUp SYBR green master mix (Thermo Fisher, A25778) and 0.5 µmol l^{-1} primers. Relative expression values were normalized to a control gene (*Rplp0*) and expressed in terms of linear relative mRNA values.

CDC permeabilization live-cell imaging. BMDMs and neutrophils were seeded at 5×10^4 cells per well on 96-well plates (E&K Scientific, EK-25090, Greiner) and BMDMs were left for 2 d before stimulation. Cells were pretreated with TLR ligands for 2–24 h depending on the experiment. For live-cell imaging, culture media were replaced by 37 °C PBS with 0.05% BSA, 1 µg ml^{-1} PI (VWR, 80057-368) and 5 µg ml^{-1} Hoechst 33342 (Thermo Fisher, H3570). Next, cells were challenged by spiking recombinant PFO (final concentration of 1 nM), gamma-irradiated SLO (final concentration of 22 nM) or recombinant ALO (final concentration of 1.5 nM), and plates were imaged every 10 min on a Molecular Devices ImageXpress XL using a ×20 objective (Nikon Plan Fluor; NA = 0.3). SLO was activated with the addition of 10 mM dithiothreitol (DTT) at 22 °C for 10 min. Total (Hoechst 33342$^+$) and permeabilized (PI$^+$) cell numbers were assessed using MetaXpress Software with Powercore using the multi-wavelength cell scoring module.

Macrophage phagocytosis assay. BMDMs were seeded at 1.5×10^5 cells per well on 24-well plates and left for 1–3 d before stimulation. Cells were pretreated with TLR ligands for 24 h. PFO was diluted in PBS to a final concentration of 2 nM, and cells were left to incubate for 15 min at 37 °C, after which cells were washed with PBS twice and the medium was replaced with fresh BMDM medium with 5% FBS. pHrodo red *S. aureus* Bioparticles Conjugate (0.4 mg ml^{-1}; Thermo Fisher, A10010) was added on top of the macrophages, and the plates were incubated for 1 h at 37 °C. To remove the non-phagocytosed apoptotic thymocytes or pHrodo red bioparticle conjugate, cells were washed once with PBS and then subjected to flow cytometry analysis.

Flow cytometry analysis. BMDMs were lifted off the plates by scraping with the plunger of a 1-ml syringe and transferred to a 96-well round-bottom plate. After pelleting, cells were first incubated with TruStain FcX (anti-mouse CD16/32) antibody (BioLegend, 101319; 1:500) in FACS buffer (PBS with 2% FBS and 1 mM EDTA) for 10 min on ice. Cells were then stained with antibodies for surface marker CD11b (BioLegend, 101207; 1:400) for 20 min on ice. After one wash, cells were resuspended in FACS buffer with 1 µg ml^{-1} DAPI (Thermo Fisher, D1306) and analyzed on a flow cytometer (Attune NxT). Data were analyzed by FlowJo V10.

Plasma membrane purification. BMDMs were seeded at 8×10^6 cells per plate on 10-cm tissue-culture-treated dishes. Cells were left for 2 d before stimulation. Cells were then treated with IFNs for 24 h. Plasma membrane purification was carried out by cell surface biotinylation followed by streptavidin affinity chromatography as previously described[36,46].

Filipin III staining. BMDMs were seeded at 1.5×10^5 cells per well on 24-well plates. Cells were left for 1–3 d before stimulation. Cells were pretreated with TLR ligands for 24 h. Filipin III was purchased from Sigma-Aldrich (F4767) and resuspended in ethanol to 2 mg ml^{-1}. BMDMs were washed once with PBS and fixed with 3% paraformaldehyde for 15 min at 22 °C. Cells were then washed twice and stained with 100 µg ml^{-1} filipin III for 30 min at 22 °C in darkness. After staining, cells were washed once with PBS and visualized using a Zeiss Axio Observer Z1.

CTB staining. BMDMs were seeded at 1.5×10^5 cells per well on 24-well plates and left for 1–3 d before stimulation. Cells were pretreated with IFNs for 24 h. BMDMs were lifted off the plates by scraping with the plunger of a 1-ml syringe and transferred to a 96-well round-bottom plate. CTB was purchased from Thermo Fisher Scientific (C34777). Cells were stained with 1 µg ml^{-1} CTB for 10 min on ice. Cells were then resuspended in FACS buffer and analyzed with a flow cytometer (Attune NxT). Data were analyzed by FlowJo V10.

Preparation of ^{15}N- or ^{13}C-labeled His-tagged ALO-D4. A plasmid for ALO-D4 (ALO amino acids 404–512 with C472A and S404C substitutions) was originally obtained from A. Radhakrishnan[36,38] (University of Texas Southwestern Medical Center, Dallas), and ^{15}N-labeled ALO-D4 was prepared as described before[56]. Briefly, ALO-D4 was expressed in BL21(DE3)pLysS *Escherichia coli* (Invitrogen) induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) in 1 liter of minimal medium containing 20.2 mM NH_4Cl and 2 g of glucose at 18 °C for 16 h. [^{15}N]NH_4Cl or [^{13}C]glucose was used for preparation of [^{15}N]ALO-D4- or [^{13}C]-ALO-D4, respectively. Cells were pelleted and lysed by sonication, and the lysate was centrifuged at 4 °C. The supernatant was mixed with 4 ml of HisPur Cobalt resin (50% bed volume; Thermo Fisher Scientific). The mixture was loaded into a column and allowed to flow through by gravity. The column was washed, and [^{15}N]ALO-D4 was eluted with a buffer containing 300 mM imidazole. The eluates were pooled and concentrated to 1 ml with an Amicon 10-kDa-cutoff concentrator (Millipore). The purified [^{15}N]ALO-D4 was stored at 4 °C.

Preparation of His-tagged OlyA and full-length ALO. Plasmids for OlyA (with C62S, C94S and S151C substitution) and ALO were kind gifts from A. Radhakrishnan[38] (University of Texas Southwestern Medical Center, Dallas).

Preparation of fluorescent ALO-D4 and OlyA. Purified protein was conjugated as described previously, with slight modifications[10]. Briefly, cysteine-substituted

ALO-D4 or OlyA was incubated with AlexaFluor-488 or AlexaFluor-594 (Life Technologies) at 4 °C overnight in 50 mM Tris-HCl (pH 7.5) and 150 mM NaCl (1× TBS) containing 1 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP). Free dye was separated by extensive buffer exchange with 1× TBS in Amicon 10-kDa-cutoff concentrator (Millipore) and stored at 4 °C.

Binding of ALO-D4 and OlyA to cells. Cells were seeded at 1.5×10^5 cells per well on poly(D-lysine)-coated coverslips (Neuvitro, Fisher Scientific, NC0672873) on a 24-well plate. After stimulation, cells were incubated with ALO-D4 or OlyA as described previously[40]. Briefly, macrophages were washed three times for 10 min in PBS with Ca^{2+} and Mg^{2+} containing 0.2% (wt/vol) BSA. Cells were then incubated with ALO-D4 or OlyA (all 20 μg ml^{-1}) in PBS with Ca^{2+} and Mg^{2+} containing 0.2% (wt/vol) BSA for 2 h at 4 °C. The unbound proteins were removed by washing three times with PBS with Ca^{2+} and Mg^{2+} for 2 min each.

Confocal microscopy analysis. After fluorescently labeled ALO-D4 or OlyA binding, cells were fixed with 3% paraformaldehyde for 15 min and stained with 5 μg ml^{-1} DAPI, and cells were washed twice with PBS with Ca^{2+} and Mg^{2+}. Images were taken using an Axiovert 200 M microscope and processed with Zen 2010 software (Zeiss). High-resolution images were captured using a Zeiss LSM880.

Quantification of fluorescence intensity. For signal quantification, cells were plated on 24-well plates (Cellvis, P24-0-N) and treated and stained as above, and fixed cells in DPBS with Ca^{2+} and Mg^{2+} were imaged within 24 h. Images were taken at the UCLA Molecular Screening Shared Resource core facility on a Molecular Devices ImageXpress confocal using a ×20 objective (Nikon Plan Fluor; NA = 0.3) on a Molecular Devices ImageXpress XL imaging system. Maximum projected cellular fluorescence intensity was assessed by MetaXpress software with Powercore using the multi-wavelength cell scoring module. Integrated fluorescence intensity profiles were exported and analyzed in R with the ggplot2 package or in GraphPad Prism 8. Filipin III staining was quantified with FIJI by measuring fluorescent signals on plasma membranes.

NanoSIMS sample preparation. As described previously[56], after ALO-D4 or OlyA binding, cells were fixed with 4% paraformaldehyde (Electron Microscopy Sciences) and 2.5% glutaraldehyde (Electron Microscopy Sciences) in 0.1 M phosphate buffer (1.14 g NaH_2PO_4 and 1.69 g Na_2HPO_4 in a 100-ml final volume of ddH_2O, pH 7.4) for 20 min at 4 °C followed by 1 h at 22 °C. The samples were washed three times for 7 min each in 0.1 M phosphate buffer, postfixed with 1% osmium tetroxide (Electron Microscopy Sciences) in 0.1 M phosphate buffer for 45 min and washed three times for 7 min each in ice-cold ddH_2O. After washing, cells were air-dried.

NanoSIMS analyses. As described before[40,56], platinum-coated (5-nm) cells were analyzed with a NanoSIMS 50l instrument (CAMECA) with some modifications. Briefly, samples were bombarded with a focused $^{133}Cs^+$ primary beam, and secondary ions (for example, $^{12}C^-$, $^{13}C^-$, $^{16}O^-$, $^{12}C^{14}N^-$ and $^{12}C^{15}N^-$) and secondary electrons were collected. Before imaging, a high $^{133}Cs^+$ primary beam (current of 1 nA; primary aperture D1 = 1) was used to presputter an area of 50×50 μm^2 for 25 s to remove the platinum coating and implant $^{133}Cs^+$. In the same region, low-magnification images (~ 40×40 μm^2) were obtained with a ~2.5-pA beam current (primary aperture D1 = 2), a dwell time of 2.5 ms per pixel and scans of 512×512 pixels. High-magnification images (~10×10 μm^2) were obtained with a ~0.8-pA beam current (primary aperture D1 = 3), a dwell time of ~10 ms per pixel and scans of 512×512 pixels.

NanoSIMS quantification. To quantify $^{13}C/^{12}C$ and $^{15}N/^{14}N$ ratios in cells, we identified particles by SEM and/or $^{12}C^-$, $^{12}C^{14}N^-$, $^{16}O^-$ or secondary electron nanoSIMS images, and regions of interest in the middle of the particles were defined with the OpenMIMS plugin in ImageJ (NIH). For each image, the mean $^{15}N/^{14}N$ and $^{13}C/^{12}C$ ratios of the regions of interest were measured by ImageJ and processed by Prism 7.0.

In vivo CDC injection. For SLO challenge studies, litter-matched control and *Ch25h*$^{-/-}$ C57BL/6J mice were shaved on the lower back and recovered overnight. On the next day, mice were intradermally injected with either PBS or SLO (8 kU per mouse). SLO was activated with 10 mM DTT at 22 °C. Photographs of the skin lesions were taken 48 h later, and the lesion area (in cm^2) was measured using ImageJ software, followed by an unequal *t*-test. For ALO challenge studies, wild-type C57BL/6 mice were shaved on the lower back and preinjected intradermally with 25HC (10 mg kg^{-1}) or ethanol vehicle control for 6 h, followed by ALO injection (20 nM). Lesion sizes were quantified as above 48 h after ALO injection. Photographs of the skin lesions were taken 48 h after the ALO challenge, and the lesion area was measured using ImageJ software, followed by an unequal *t*-test. Samples were collected for histologic analysis in the UCLA Translational and Pathology Core Laboratory.

Statistical analysis. Experiments were conducted as three biologically distinct independent replicates unless otherwise stated. Statistical significance was determined using GraphPad Prism software. For bar graphs, line graphs and scatterplots, data are shown as the mean and the error bars depict the s.e.m. unless otherwise indicated. For all CDC-permeabilization assays, an average of 800 cells were imaged per well and used for calculations. Statistical analysis and assessment of CDC-permeabilization curves began when the first images were taken after CDCs were added to cultures until the end of the assay at 60 min. Statistical tests used for each graph are specified in the figure legend.

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

Data availability

All original data are available from the corresponding author upon request.

References

52. Divakaruni, A. S. et al. Thiazolidinediones are acute, specific inhibitors of the mitochondrial pyruvate carrier. *Proc. Natl Acad. Sci. USA* **110**, 5422–5427 (2013).
53. Bligh, E. G. & Dyer, W. J. A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **37**, 911–917 (1959).
54. Williams, K. J. et al. An essential requirement for the SCAP/SREBP signaling axis to protect cancer cells from lipotoxicity. *Cancer Res.* **73**, 2850–2862 (2013).
55. Argus, J. P. et al. Development and application of FASA, a model for quantifying fatty acid metabolism using stable isotope labeling. *Cell Rep.* **25**, 2919–2934 (2018).
56. He, C., Fong, L. G., Young, S. G. & Jiang, H. NanoSIMS imaging: an approach for visualizing and quantifying lipids in cells and tissues. *J. Investig. Med.* **65**, 669–672 (2017).

Acknowledgements

This research was supported by NIH grants AI093768 (to S.J.B.), HL146358 (to S.J.B. and P.T.), AR073940 (to P.O.S), and HL136543 (to E.J.T.). M.S.L. is supported by Ruth L. Kirschstein National Research Service Award AI007323. The research described was also supported by an NIH/National Center for Advancing Translational Science (NCATS) UCLA CTSI grant (UL1TR001881). We thank S. Young, T. Weston and R.S. Jung for help with protein purification. We thank T. Weston for NanoSIMS sample preparation and SEM imaging. We thank A. Divakaruni for guidance with PFO permeabilization assays. We thank A. Radhakrishnan for ALO-D4 and full-length ALO plasmids. We thank S. Young, A. Hoffmann, Y. Du, R. Sun, T.-T. Wu, J. F. Miller and M. Li for thoughtful discussions.

Author contributions

S.J.B. conceived the study. Q.D.Z. led the design and execution of experiments. X.C. codesigned and performed all flow cytometry experiments and data analysis. V.L.B., W.Y.H. and J.J.M. contributed to flow cytometry experiments. J.J.M. and M.S.L. contributed to protein purification and staining experiments. Q.D.Z., M.S.L. and R.D. developed and performed live-cell imaging assays. C.H. performed NanoSIMS analysis and contributed to protein purification. J.J.M. and E.B.K. contributed to RNA analysis. W.Y.H. and A.G.Y. performed GC–MS analysis with help from Q.D.Z. and E.B.K. K.J.W. conducted lipidomic studies. X.X., A.F., P.T. and E.J.T. contributed to *Abcg1*, *Abca1* and SCAP KO studies. A.E.D. analyzed gene expression data. A.-C.F., P.O.S., M.S.L. and S.T.S. conceptualized and developed the in vivo SLO challenge assay. W.Y.H. contributed to data visualization. S.J.B., Q.D.Z., W.Y.H., X.C. and P.T. contributed to construction of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

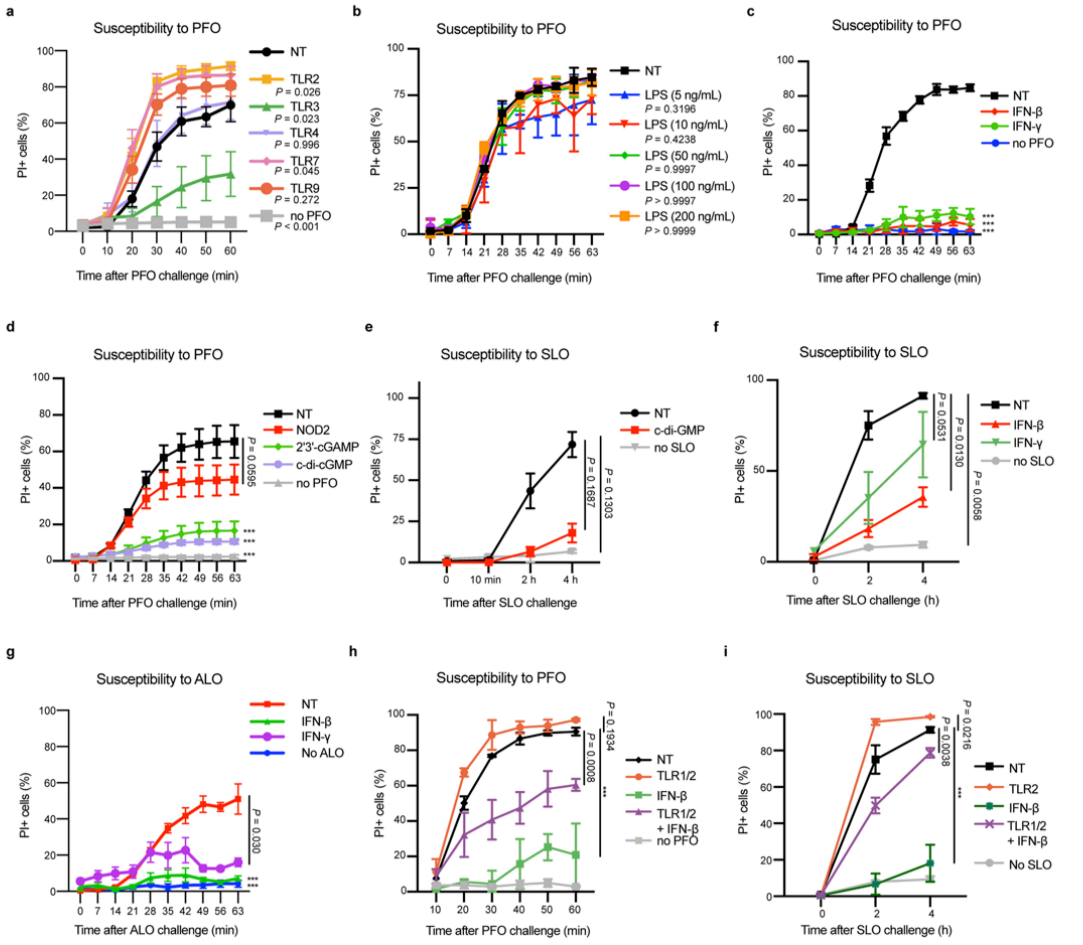
Extended data is available for this paper at https://doi.org/10.1038/s41590-020-0695-4.

Supplementary information is available for this paper at https://doi.org/10.1038/s41590-020-0695-4.

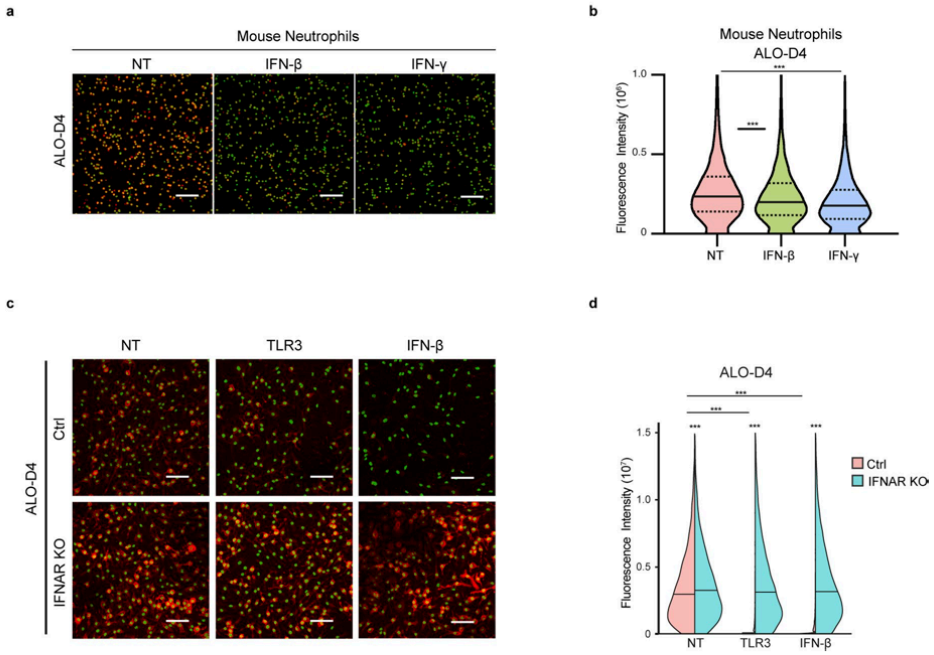
Correspondence and requests for materials should be addressed to S.J.B.

Reprints and permissions information is available at www.nature.com/reprints.

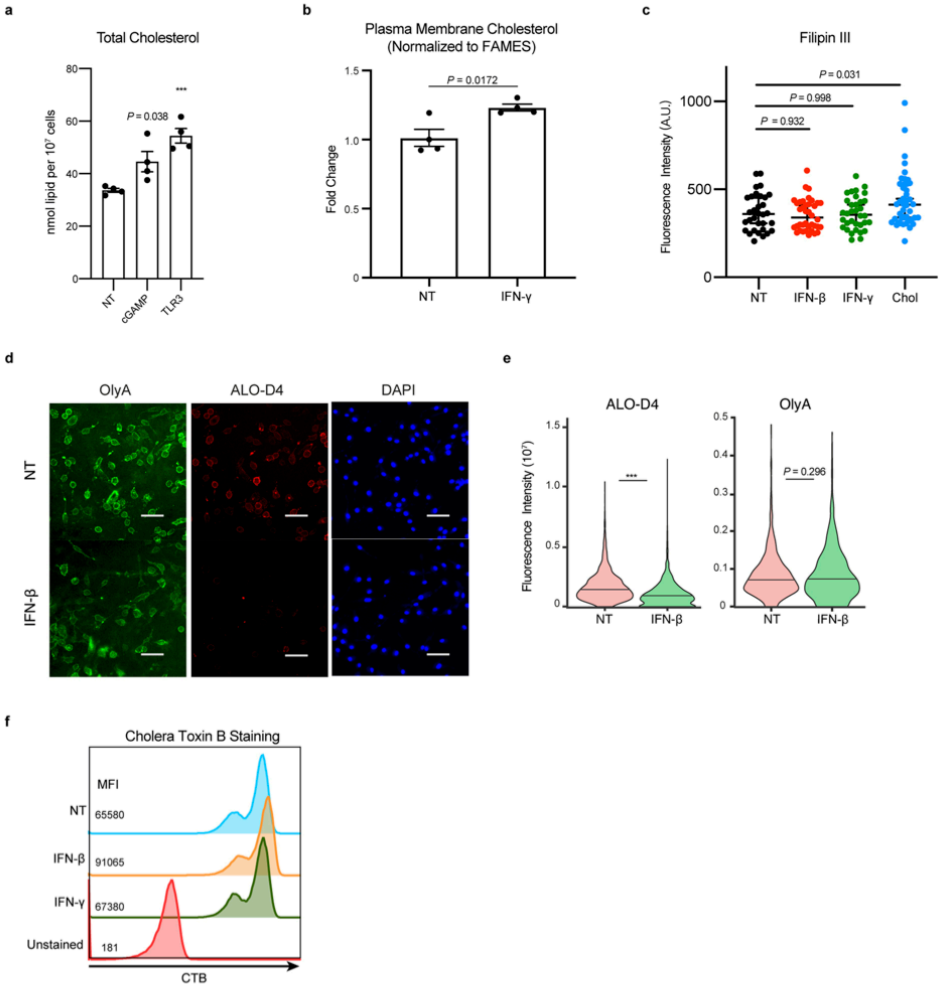
Editor recognition statement L. A. Dempsey was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.



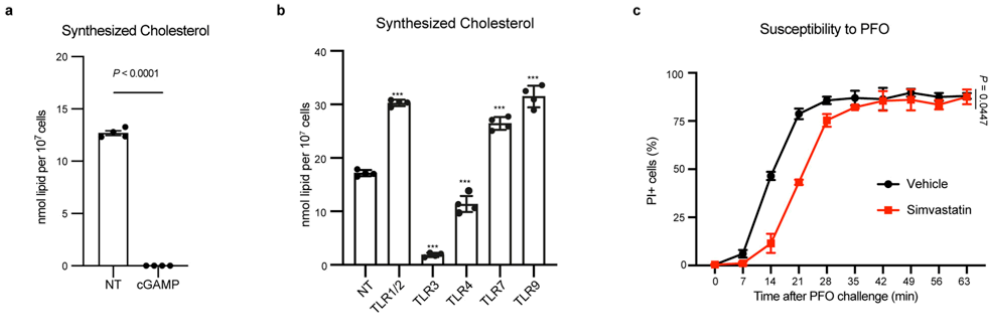
Extended Data Fig. 1 | Interferon signaling mediates resistance to cholesterol-dependent cytolysins. a, Percentage of PI-positive BMDMs treated with TLR1/2 agonist (Pam3CSK4; 50 ng/mL), TLR3 agonist (Poly(I:C); 1 μg/mL), TLR4 agonist (LPS; 50 ng/mL), TLR7 agonist (CL307; 100 nM), TLR9 (ODN1668; 100 nM) agonist, or unstimulated (NT) for 24 h and then challenged with PFO for up to 60 min in the presence of PI. Cells were imaged every 10 min to assess changes in PI incorporation. **b**, Percentage of PI-positive BMDMs treated with the various concentrations of TLR4 agonist for 24 h and then challenged with PFO for up to 60 min in the presence of PI. Cells were imaged every 7 min to assess changes in PI incorporation. **c**, Percentage of PI-positive BMDMs treated with IFN-β or IFN-γ (20 ng/mL) for 24 h and then challenged with PFO for 60 min in the presence of PI. **d**, Percentage of PI-positive BMDMs treated with NOD2 agonist (N-Glycolyl-MDP; 20 μg/ml), STING agonist (2',3'cGAMP and c-di-GMP; both 2 μg/mL) for 24 h and then challenged with PFO for 60 min in the presence of PI. **e**, Percentage of PI-positive BMDMs treated with STING ligand for 24 h and then challenged with Streptolysin O (SLO) for 4 h in the presence of PI. **f**, Percentage of PI-positive BMDMs treated with IFN-β or IFN-γ (20 ng/mL) for 24 h and then challenged with SLO for 4 h in the presence of PI. **g**, Percentage of PI-positive BMDMs treated with IFN-β or IFN-γ (20 ng/mL) for 24 h and then challenged with ALO for 60 min in the presence of PI. **h**, Percentage of PI-positive BMDMs treated with IFN-β (100 ng/mL) or TLR1/2 agonist, or IFN-β (100 ng/mL) together with TLR1/2 agonist for 24 h and then challenged with PFO for 60 min in the presence of PI. **i**, Percentage of PI-positive BMDMs treated with IFN-β (100 ng/mL) or TLR1/2 agonist, or IFN-β (100 ng/mL) together with TLR1/2 agonist for 24 h and then challenged with SLO for 4 h in the presence of PI. Data are representatives of three independent experiments, and are shown as mean ± s.e.m. (n = 3). Statistical significance was determined using an RM one-way ANOVA with Dunnett's correction (**a**-**g**) or a two-way ANOVA with Dunnett's correction (**h**, **i**). ***$P<0.001$.



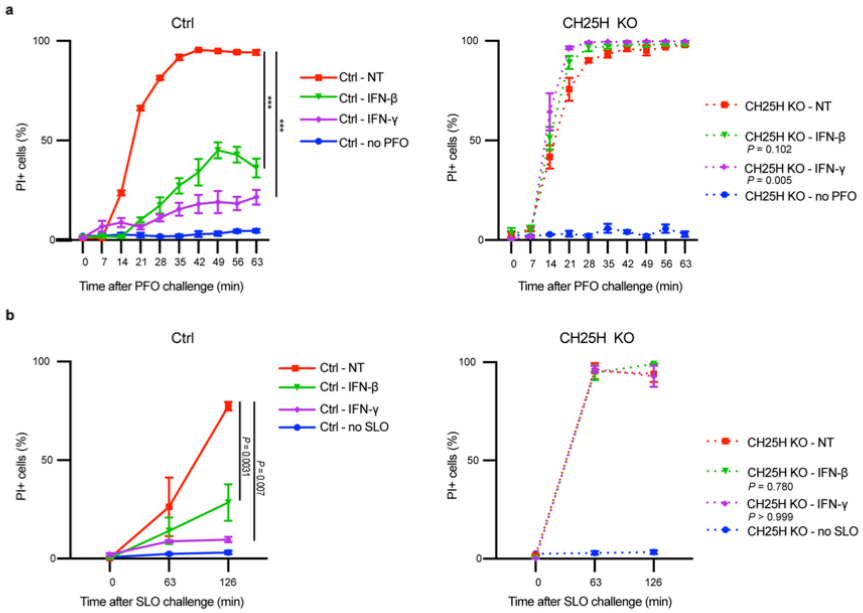
Extended Data Fig. 2 | IFN signals decrease plasma membrane binding to ALO-D4 protein. a, Confocal images of neutrophils stimulated with IFN-β or IFN-γ (20 ng/mL) for 6 h, and then stained with fluorescent ALO-D4 and DAPI. **b**, Violin plots of cellular fluorescent intensity quantified from **a** (n = 20334, 18546, 16290). **c**, Confocal images of WT or type I interferon receptor–deficient (IFNAR KO) BMDMs stimulated with TLR3 agonist (1 μg/mL) or IFN-β (20 ng/mL) for 24 h, and then stained with fluorescent ALO-D4 and DAPI. **d**, Violin plots of cellular fluorescent intensity quantified from **c** (n = 5543, 6682, 4673, 8231, 5201, 7906). Data are representatives of three independent experiments. Violin plots are shown with median (solid lines in **b**, **d**) and 25% and 75% percentiles (dashed lines in **b**), and statistical significance was determined using a Kruskal-Wallis test with Dunn's correction. ***P<0.001. Scale bars in **a**, **c** represent 50 μm.



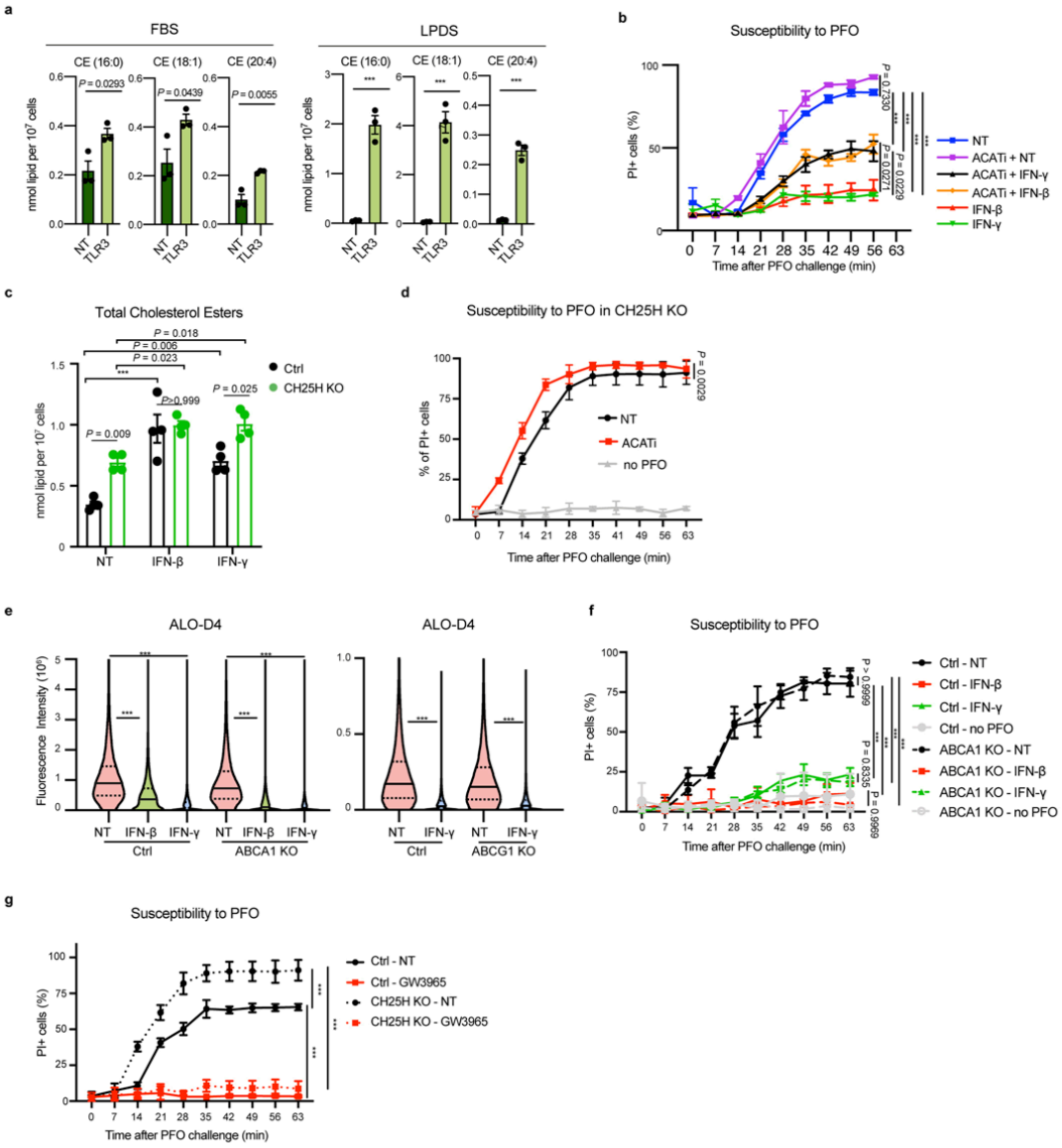
Extended Data Fig. 3 | IFN signals reprogram cholesterol metabolism to decrease the pool of cholesterol targeted by CDCs. a, Total cholesterol (nmol/10^7 cells) from C57BL/6 bone marrow-derived macrophages (BMDMs) stimulated with cGAMP (2 μg/mL), TLR3 agonist (Poly(I:C); 1 μg/mL), or unstimulated (NT) for 48 h. Total cholesterol was determined by GC-MS (n = 4). **b**, Total plasma membrane cholesterol (normalized to total FAMEs) from C57BL/6 bone marrow-derived macrophages (BMDMs) stimulated with IFN-γ (40 ng/mL) or unstimulated (NT) for 24 h (n = 4). **c**, Relative Fillipin III fluorescence intensity of plasma membranes of untreated macrophages or macrophages stimulated with IFN-β (20 ng/mL) or IFN-γ (20 ng/mL) for 24 h (n = 32, 38, 34, 42). MβCD-Cholesterol loaded macrophages indicate dynamic range of Fillipin III fluorescence and are included as a positive control. **d**, Confocal images of BMDM stimulated with IFN-β (20 ng/mL) for 24 h, and then stained with fluorescent ALO-D4 or OlyA and DAPI. Scale bar, 50 μm. **e**, Violin plots of cellular fluorescent intensity quantified from **d** (n = 2225, 2021, 2225, 2021). **f**, Cholera Toxin B staining of BMDM stimulated with IFN-β or IFN-γ (20 ng/mL) for 24 h. Median fluorescence intensity (MFI) are indicated on the left. Data are representative of three (**a**, **d**, **e**, **f**) independent experiments, three independent samples (**c**) or from 4 biological replicates (**b**). Data in **a**-**c** are shown as mean ± s.e.m., violin plots in **e** are shown with median (solid lines). Statistical significance was determined using an unpaired two-tailed Student's *t*-test (**a**, **b**), a one-way ANOVA with Dunnett's correction (**c**), or a two-tailed Mann-Whitney test (**e**) ***$P<0.001$.



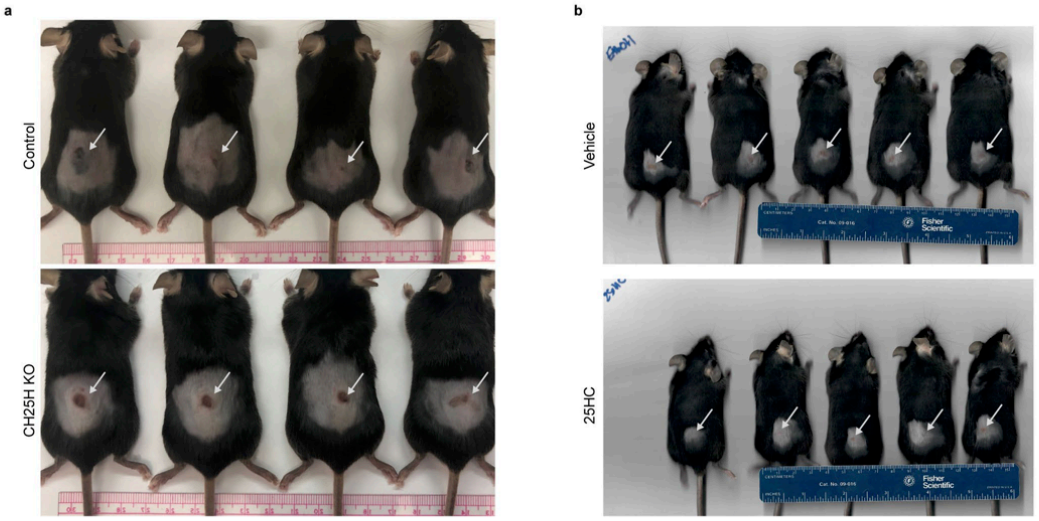
Extended Data Fig. 4 | Cholesterol synthesis is linked to CDC susceptibility. a, Net synthesized cholesterol (nmol/10^7 cells) from C57BL/6 bone marrow-derived macrophages (BMDMs) stimulated with cGAMP (2 μg/mL), or unstimulated (NT) for 48 h. Synthesized cholesterol was determined by GC-MS and isotopomer spectral analysis modeling (n = 4). **b**, Net synthesized cholesterol (nmol/10^7 cells) from C57BL/6 bone marrow-derived macrophages (BMDMs) stimulated with TLR1/2 agonist (Pam3CSK4; 50 ng/mL), TLR3 agonist (Poly(I:C); 1 μg/mL), TLR4 agonist (LPS; 50 ng/mL), TLR7 agonist (CL307; 100 nM), TLR9 (ODN1668; 100 nM) agonist, or unstimulated (NT) for 48 h. Synthesized cholesterol was determined by GC-MS and isotopomer spectral analysis modeling (n = 4). **c**, Percentage of PI-positive WT BMDMs treated with Simvastatin (1 μM) for 4 h and then challenged with PFO for 60 min in the presence of PI (n = 3). Data are representative of three independent experiments and are shown as mean + s.e.m. Statistical significance was determined using an unpaired two-tailed Student's *t*-test (**a**), a one-way ANOVA with Dunnett's correction (**b**), or a paired two-tailed Student's *t*-test (**c**). ***P<0.001.



Extended Data Fig. 5 | Production of 25-hydroxycholesterol is required to maintain changes in plasma membrane cholesterol and mediates resistance to CDCs. a, Percentage of PI-positive control or CH25H KO BMDMs stimulated with IFNs (20 ng/mL) for 24 h and then challenged with PFO for 60 minutes in the presence of PI. **b**, Percentage of PI-positive control or CH25H KO BMDMs stimulated with IFNs (20 ng/mL) for 24 h and then challenged with SLO for 2 h in the presence of PI. Data are representative of three independent experiments and are shown as mean + s.e.m. (n = 3) and statistical significance was determined using a two-way ANOVA with Dunnett's correction. ***$P<0.001$.



Extended Data Fig. 6 | Cholesterol esterification contributes to CDC resistance of macrophages. a, Quantification (nmol/10^7 cells) of cholesterol ester species (16:0. 18:1, 20:4) in BMDMs stimulated with TLR3 agonist (1 μg/mL) in FBS or LPDS for 48 h. CE species pool sizes were determined by direct infusion MS. **b**, Percentage of PI-positive WT BMDMs treated with IFN-β, or IFN-γ (20 ng/mL), or in combination with ACATi 58-035 (4.3 μM) for 24 h and then challenged with PFO for 60 min in the presence of PI. **c**, Quantification (nmol/10^7 cells) of total cholesterol ester (CE) in control or CH25H KO BMDMs stimulated with IFN-β (20 ng/mL) or IFN-γ (20 ng/mL) for 48 h. CE pool sizes were determined by direct infusion mass spectrometry. **d**, Percentage of PI-positive CH25H KO BMDMs treated with ACATi 58-035 (4.3 μM) for 24 h and then challenged PFO for 60 min in the presence of PI. **e**, Violin plots of cellular fluorescent intensity quantified from control or ABCA1 KO or ABCG1 KO BMDMs stimulated with IFNs (20 ng/mL) for 24 h and then stained with fluorescent ALO-D4 and DAPI (n = 5943, 4126, 5727, 6914, 5740, 7898; n=7201, 7532, 7563, 7417). **f**, Percentage of PI-positive control or ABCA1 KO BMDMs treated with IFNs (20 ng/mL) for 24 h and then challenged PFO for 60 min in the presence of PI. **g**, Percentage of PI-positive control or CH25H KO BMDMs treated with LXR agonist GW3965 (1 μM) for 24 h and then challenged PFO for 60 min in the presence of PI. Data are representatives of two (**a**, **c**) or three (**b**, **d**, **e**, **f**, **g**) independent experiments. Data in **a**, **b**, **c**, **d**, **f** and **g** are shown as mean + s.e.m. (n = 3 in **a**, **b**, **d**, **g**; n = 4 in **c**). Violin plots in **e** are shown with median (solid lines) and 25% and 75% percentiles (dashed lines). Statistical significance was determined using an unpaired two-tailed Student's *t*-test (**a**), a two-way ANOVA with Tukey's correction (**b**, **c**, **f**, **g**), a paired two-tailed Student's *t*-test (**d**), or a Kruskal–Wallis test with Dunn's correction (**e**). ***$P<0.001$.



Extended Data Fig. 7 | 25HC mediates protection to CDC induced tissue damage. a, Lesion images of control or CH25H KO mice challenged intradermally with SLO (8 kU/mouse) for 48 h. **b**, Lesion images of vehicle or 25HC pretreated mice challenged intradermally with ALO (20 nM) for 48 h.

APPENDIX B

NF-κB Broadly Orchestrates Nucleosome Remodeling during the Primary Response to Toll-Like Receptor 4 Signaling

An-Chieh Feng,[1,2,6] Brandon J. Thomas,[1,2,6] Xin Liu,[1,2] Allison E. Daly,[1,2] Fei Sun,[2,3] Lijing Cheng,[4] Michael F. Carey,[2,3] Philip O. Scumpia,[2,4] and Stephen T. Smale[1,2,4,5,7*]

[1]Department of Microbiology, Immunology, and Molecular Genetics, [2]Molecular Biology Institute, [3]Department of Biological Chemistry, [4]Department of Medicine, and [5]Howard Hughes Medical Institute
University of California, Los Angeles, Los Angeles, CA 90095, USA
[6]These authors contributed equally
[7]Lead contact
[*]Correspondence: smale@mednet.ucla.edu

SUMMARY

Inducible nucleosome remodeling at hundreds of latent enhancers and several promoters helps shape the transcriptional response to Toll-like receptor 4 (TLR4) signaling in macrophages. However, the identities of the transcription factors that promote TLR-induced remodeling have remained elusive. Analysis strategies that enrich ATAC-seq profiles for genomic regions most likely to undergo remodeling initially revealed a unique relationship between NF-κB and TLR4-induced remodeling events. A critical functional role for NF-κB in remodeling was then revealed by CRISPR-Cas9 mutagenesis of NF-κB genes and binding motifs. This critical role was found to be broad and possibly universal during the TLR4 primary response. Remodeling selectivity at defined regions is conferred by collaboration between NF-κB and other inducible factors, including IRF3 and MAP kinase-induced factors. Thus, NF-κB is unique among TLR4-induced transcription factors in its broad contribution to inducible nucleosome remodeling, alongside its well-established ability to activate poised enhancers and promoters assembled into open chromatin.

Keywords: macrophages, transcription, chromatin, nucleosome remodeling, NF-κB, IRF-3

INTRODUCTION

The rapid and selective activation of hundreds of genes is required for innate immune cells to mount a tailored response to each of a broad range of microbial and environmental stimuli. The selectivity of the transcriptional response is dictated by the cell-surface and intracellular sensors activated by the stimulus, by the downstream signaling pathways, by the transcriptional and chromatin regulators activated by these pathways, and by signaling dynamics and kinetics (Heinz et al., 2015; Glass and Natoli, 2016; Monticelli and Natoli, 2017; Natoli and Ostuni, 2019; Sheu and Hoffmann, 2022).

The differential responses to Toll-like receptor 4 (TLR4) and TLR2, two pattern recognition receptors activated by distinct microbial products, serve as a classic example of response selectivity. TLR4 strongly activates both the MyD88 and TRIF signaling pathways, which in turn activate a defined set of transcription factors and their downstream target genes. In contrast, TLR2 potently activates only MyD88 and therefore fails to activate a key set of TLR4-induced genes (Toshchakov et al., 2002; Yamamoto et al., 2003; Hoebe et al., 2003). Transcription factors induced solely by post-translational mechanisms regulate the primary response to the stimulus; additional factors encoded by primary response genes help shape the subsequent secondary response (Ramirez-Carrozzi et al., 2006, 2009; Hargreaves et al., 2009).

Chromatin collaborates with signaling pathways and transcription factors to orchestrate stimulus-specific transcriptional responses via two fundamentally distinct mechanisms (Gatchalian et al., 2020; Martínez de Paz and Josefowicz, 2021). First, the chromatin landscape assembled during the development of a responsive cell type dictates which control regions and genes are capable of participating in the response to a stimulus. In mature unstimulated macrophages, the chromatin landscape is characterized by thousands of poised enhancers that exhibit physical accessibility and histone modifications (e.g. histone H3K4me1) associated with poised but inactive transcriptional states (Gisletti et al., 2010; Heinz et al., 2010; Barozzi et al., 2014). The promoters of almost all TLR-induced genes are also assembled into accessible chromatin in unstimulated macrophages. NF-κB and AP-1 are among a collection of inducible

transcription factors that frequently bind both accessible promoters and poised enhancers to activate transcription (Gisletti et al., 2010; Heinz et al., 2010; Tong et al., 2016).

Chromatin also contributes directly to the inducibility and selectivity of the transcriptional response (Gatchalian et al., 2020; Martínez de Paz and Josefowicz, 2021). Thousands of latent enhancers remain physically inaccessible and lack histone modifications characteristic of a poised or active state until the cells encounter a stimulus (Kaikkonen et al., 2013; Otsuni et al., 2013). A small number of TLR4-induced promoters also undergo stimulus-responsive nucleosome remodeling to enhance accessibility, with remodeling playing a critical role in response selectivity (Weinmann et al., 1999; Ramirez-Carrozzi et al., 2009; Tong et al., 2016; Comoglio et al., 2019). In one example, *Ccl5* was the only TLR4-induced primary response gene to exhibit a strong requirement for the transcription factor IRF3 for inducible remodeling at its promoter (Tong et al., 2016). This finding suggests a remarkable degree of selectivity in which promoter nucleosomes help prevent expression of the Ccl5 chemokine following exposure to many stimuli, with induction limited to those stimuli that induce IRF3 (Tong et al., 2016).

Stimulus-responsive nucleosome modeling is typically catalyzed by ATP-dependent nucleosome remodeling complexes, including BAF (SWI/SNF) complexes, which can promote nucleosome eviction, sliding, or conformational transitions (Hargreaves and Crabtree, 2011; Clapier et al., 2017). However, a key unresolved issue in the field is the identities of the transcription factors that promote TLR-induced remodeling. Comoglio et al. (2019) provided evidence of combinatorial regulation of remodeling by multiple inducible factors through a detailed analysis of TLR-induced chromatin changes combined with modeling and ChIP-seq analyses. A predictive computational tool used in this study identified NF-κB as the strongest candidate to induce remodeling (Comoglio et al., 2019), despite the frequent binding of NF-κB to open chromatin (Gisletti et al., 2010; Heinz et al., 2010; Tong et al., 2016). Further evidence that NF-κB may participate in nucleosome remodeling emerged from a study in which IκBα deficiency was used to prolong NF-κB activation in response to tumor necrosis factor α stimulation, leading to increased accessibility at a cluster of control regions that were unaffected by transient NF-κB

activation (Cheng et al., 2021). Transcription factors of the TLR-inducible AP-1 family have also been implicated in the regulation of nucleosome remodeling, but primarily in TLR-independent settings (Vierbuchen et al., 2017; Gao et al., 2019; Wolf et al., 2023).

A challenge in uncovering the mechanisms regulating inducible remodeling is that the genomic locations of inducible remodeling events are difficult to define at a genome-wide scale. Assays like ATAC-seq and nuclease hypersensitivity-seq reveal genomic regions exhibiting physical changes in DNA accessibility (Buenrostro et al., 2015; Comoglio et al., 2019). However, it is unlikely that most statistically significant physical changes represent true nucleosome remodeling events, as changes in transcription factor-DNA interactions at open chromatin regions can also alter accessibility (Buenrostro et al., 2013; Li et al., 2019). Another challenge is that BAF nucleosome remodeling complexes bind to most open chromatin regions genome-wide (Hargreaves, 2021), making it difficult to use BAF ChIP-seq profiles to definitively identify locations of inducible remodeling. Moreover, BAF complexes are broadly required for the maintenance of open chromatin and their absence results in cell death over time (Ramirez-Carrozzi et al., 2009; Hargreaves, 2021), limiting the utility of loss-of-function experiments.

To gain further insight into the regulation of TLR4-induced nucleosome remodeling, we initially used ATAC-seq profiling in lipid A-stimulated macrophages. However, we concentrated on genomic regions exhibiting the largest and most consistent changes in accessibility rather than employing typical analysis approaches focused on all regions that exhibit statistically significant increases. This approach, followed by extensive functional analyses, provided strong evidence that NF-κB plays a surprisingly broad and possibly universal role in remodeling during the primary TLR4 response. Further genomic and functional studies revealed that remodeling selectivity during the primary response is conferred by functional collaborations at distinct subsets of regulatory regions between NF-κB and other TLR-induced transcription factors, including IRF3 and MAP kinase (MAPK)-induced factors.

RESULTS

ATAC-seq Profiling in Response to TLR4 Signaling

ATAC-seq was performed with mouse bone marrow-derived macrophages (BMDMs) stimulated with the TLR4 ligand, lipid A, for 0, 30, 60, or 120 min to identify in an unbiased manner regions of the mouse macrophage genome that exhibit changes in chromatin accessibility. Merger of reads from multiple biological replicates revealed 101,448 accessible DNA regions (i.e. statistically significant sequencing read peaks) that were reproducibly observed in at least one time point. By analyzing ATAC-seq reads per kbp of DNA per million mapped reads (RPKM), most peaks (72.8%) increased or decreased in response to the stimulus by less than 2-fold among all stimulated time points in comparison to unstimulated cells (Figure B-1A), with 2.7% of peaks exhibiting repression of more than 2-fold as their maximum change. In contrast, 19.4% were maximally induced 2-5-fold, with 5.1% (5,178 peaks) induced more than 5-fold. Most ATAC-seq peaks were located in intronic and intergenic regions, with 10.1% located in promoters (between -1kb and +100 bp relative to a TSS; Figure B-1B). However, peaks induced >5-fold were scarce within promoters, consistent with prior evidence that inducible promoters rarely exhibit inducible remodeling (Figure B-1B, Tong et al. 2016). Examples of Integrated Genomics Viewer (IGV) tracks (Robinson et al., 2011) for repressed, constitutive and induced promoter peaks are shown in Figure B-1C. Notably, a visual analysis revealed that several constitutive and inducible genes exhibit continuous signals throughout the transcription unit, presumably due to chromatin accessibility linked to active transcription (Figure B-1C, right).

We next divided all ATAC-seq peaks into 40 bins based on maximum fold-change in accessibility during the time course, and then examined the distribution of genomic locations in each bin (Figure B-1D). Consistent with Figure B-1B, promoters represent a large fraction of peaks in bins that exhibit little change in ATAC-seq signal following stimulation, but only a small fraction in the most induced and repressed bins. Importantly, although most peaks induced more than 2-fold are statistically significant ($p<0.05$, data not shown), a large fraction of these peaks are present in both unstimulated and stimulated cells, with the fraction of peaks observed in

unstimulated cells declining as the average fold-activation in each bin increases (Figure B-1E). For example, in Bin 30 (peaks induced approximately 2-fold), half of the genomic regions exhibited called peaks in unstimulated cells, whereas very few peaks in Bin 40 are observed prior to stimulation.

The finding that large numbers of ATAC-seq peaks with weak induction by lipid A have called peaks in unstimulated cells suggests that these genomic regions may often contain constitutively open chromatin, with the small induction of ATAC-seq signal due to smaller changes in chromatin or DNA structure associated with inducible transcription factor binding rather than nucleosome remodeling. Because of this issue, we reasoned that the most strongly induced ATAC-seq peaks, where peaks are generally undetectable in unstimulated cells and at which the largest fold inductions are observed, deserve special attention, as they represent the genomic regions that are most likely to undergo bona fide nucleosome remodeling events. (Simply eliminating locations with called peaks in unstimulated cells was undesirable because the vast majority of weakly induced peaks that remained barely exceeded the peak-calling threshold in stimulated cells and were therefore unreliable.) Therefore, for most of the analyses described below, we divided peaks into bins on the basis of their magnitude of ATAC-seq induction (stimulated/unstimulated RPKM), or we focused on ATAC-seq peaks induced >5-fold with statistical criteria added to further increase our emphasis on the genomic regions that are most likely to undergo nucleosome remodeling. The term “nucleosome remodeling” will be used below to refer to the most strongly and consistently induced ATAC-seq signals, with the caveat that we are unable to definitively establish that these large physical changes in accessibility in vivo are analogous to ATP-dependent nucleosome remodeling events defined in vitro.

Figure B-1F shows genome tracks for the TLR4-induced *Ccl9* gene as an example of the above points. Three ATAC-seq peaks, including the *Ccl9* promoter peak, show weak induction (1.2-2.2-fold) and are also present in unstimulated cells, whereas a fourth intergenic peak was induced 14-fold and lacked a called peak in unstimulated cells. Although the reason for each

change in physical accessibility remains unknown, we reasoned that the fourth genomic region is the most likely to reflect TLR4-induced nucleosome remodeling.

ChIP-seq analysis of binding by the Brg1 catalytic subunit of BAF complexes provided further insight (Figure B-1G). In unstimulated BMDMs, Brg1 peaks were detected at 20-50% of the ATAC-seq peaks that were induced by only small magnitudes in response to the stimulus (e.g. bins 3-30), with little change in these percentages following stimulation. In contrast, in bin 40, containing the most strongly induced ATAC-seq peaks, Brg1 binding in unstimulated cells was observed at only 0.8% of the ATAC-seq peaks, with Brg1 peaks detected at 37% of the ATAC-seq peaks following stimulation. This result reveals a close correlation between inducible Brg1 binding and inducible ATAC-seq signals, but only in bins containing the most strongly induced ATAC-seq peaks. Similar results were obtained with ChIP-seq datasets for histone H3K27ac (a histone modification associated with active transcription) and p300 (a co-activator that associates with active transcriptional regulatory regions), as the bins showing the strongest ATAC-seq induction also exhibited the strongest increases in H3K27ac and in p300 binding following stimulation (data not shown). Taken together, these results are consistent with the hypothesis that bins with peaks displaying the strongest ATAC-seq induction are the most enriched in genomic regions assembled into inaccessible nucleosomes in mature, unstimulated BMDMs; moreover, these regions are much more likely to undergo TLR4-induced Brg1 recruitment and nucleosome remodeling than peaks exhibiting smaller increases in ATAC-seq signals.

Characteristics of Inducible Promoter and Intergenic ATAC-seq Peaks

Our prior analysis of 132 TLR4-induced primary response genes (i.e. in general, genes induced >10-fold during a 2-hr time-course of lipid A stimulation) revealed strong ATAC-seq induction at the promoters of only three of the 132 genes: *Ccl5*, *Gbp5*, and *Irg1* (Tong et al., 2016). Using our current thresholds and data sets, induced ATAC-seq signals (>5-fold) were observed at these three promoters and four additional primary response promoters (Figure B-S1A). Induced ATAC-seq signals (>5-fold) were also observed at the promoters of six of our previously defined

secondary response genes (Figure B-S1A). Further analysis across the 2-hr time course of all promoters throughout the genome (-500 to +150) revealed inducible ATAC-seq signals (>5-fold) at only 26 additional promoters (Figure B-S1A). Twenty were associated with weakly induced genes, with the remaining six showing little or no transcriptional induction. The results confirm that strongly induced ATAC-seq signals at promoters are rare.

We next analyzed intergenic regions. As an initial test of the relevance of intergenic ATAC-seq peaks, we calculated the distance from each TLR4-induced intergenic peak to the nearest gene that exhibited transcriptional induction of at least 10-fold, as determined by chromatin-associated RNA-seq (Tong et al., 2016). This analysis revealed that strongly induced ATAC-seq peaks (>5-fold) are greatly over-represented in the vicinity of strongly induced genes; weakly induced ATAC-seq peaks are over-represented to a lesser extent, with little or no over-representation of constitutive and repressed peaks (Figure B-S1B).

K-means cluster analysis revealed that the strongly induced intergenic peaks exhibited diverse kinetic profiles (Figure B-S1C). Distinct subsets of peaks were either resistant to or sensitive to the protein synthesis inhibitor, cycloheximide (CHX), demonstrating inducible remodeling during both the CHX-resistant primary response and the CHX-sensitive secondary response (Figure B-S1C).

Transcription Factor Binding Site Motif Enrichment

To identify transcription factors that may contribute to TLR4-induced nucleosome remodeling, motif analysis was performed with the 40 bins representing all ATAC-seq peaks, as in Figure B-1. This analysis identified several transcription factor families that exhibit differential enrichment across the fold-induction range (Figure B-2A). Motifs for bZIP, ETS, IRF, NF-κB (Rel-homology domain, RHD), POU/Homeobox, Runt family members, and a small subset of zinc finger family members exhibited the greatest enrichment in the bins containing strongly induced ATAC-seq peaks (Figure B-2A). Strikingly, NF-κB motifs stood out in showing an unusually close correlation with the most strongly induced ATAC-seq peak bins in comparison to the more weakly induced

peak bins (Figure B-2B). That is, the percentage of peaks containing an NF-κB motif differed between the strongly induced (>5-fold) and weakly induced (2-5-fold) bins to a much greater extent than observed with other motifs. In particular, bZIP and IRF motifs displayed a more gradual enrichment across the ATAC-seq induction spectrum (Figure B-2B). This distinctive property of NF-κB provided initial evidence of a close and unusual relationship between NF-κB and TLR4-induced nucleosome remodeling.

To separate putative nucleosome remodeling events associated with the primary versus secondary responses to TLR4 signaling, ATAC-seq was performed in BMDMs stimulated with lipid A for 0, 30, 60, and 120 min in the presence of CHX. Strongly induced peaks (>5-fold) were placed into eight bins on the basis of the extent to which CHX altered the induced ATAC-seq signal. NF-κB motifs were over-represented most strongly in the peaks exhibiting the greatest CHX resistance (primary response), with IRF and POU/Homeobox motifs over-represented in CHX-sensitive peaks (secondary response) (Figure B-2C). The over-representation of IRF motifs during the CHX-sensitive secondary response is consistent with prior evidence that nucleosome remodeling contributes to the Type 1 interferon response, which dominates the secondary response to TLR4 (Liu et al., 2002).

Together, the above results show that NF-κB motifs are the only identifiable transcription factor binding motifs that exhibit a strong, preferential enrichment at ATAC-seq peaks that are strongly induced during the TLR4 primary response. NF-κB motifs are also enriched at weakly induced peaks (which are less likely to reflect nucleosome remodeling), but by a much smaller magnitude. The much broader enrichment of bZIP motifs across the fold-induction spectrum does not rule out a role in nucleosome remodeling, but suggests a different relationship across the spectrum (see Discussion).

Broad Binding of NF-κB to Strongly Induced ATAC-seq Peaks

To evaluate further the relationship between NF-κB and ATAC-seq peaks, we examined ChIP-seq data sets for the NF-κB family member, RelA. When examining 40 bins of peaks defined on

the basis of fold-induction, RelA binding was observed most frequently in bin 40, with RelA binding to 34% of peaks in this bin (Figure B-3A), When the 40 bins were created based on the statistical significance (p-value) of ATAC-seq induction instead of fold-induction, 59% of peaks in the highest confidence bin exhibited RelA binding (Figure B-3A).

Because low p-values are not restricted to ATAC-seq peaks exhibiting large fold-changes following induction (low p-values can be obtained with small fold-changes if the changes are highly reproducible), we focused on 3,289 primary response peaks exhibiting strong average induction (>5-fold) and then separated them into ten bins on the basis of their p-value. In this analysis, a remarkable 82.7% (272 of 329) of the primary response peaks in the highest confidence bin (bin 10) bound RelA (Figure B-3B). The increasing frequency of RelA binding across the ten bins, from 10.4% of peaks with RelA binding in bin 1 to 82.7% in bin 10, highlights the value of combining stringency fold-induction and statistical significance criteria.

The above results suggest that RelA binding to the most strongly and consistently induced ATAC-seq peaks is extremely common, but raises the question of whether the 17.3% of ATAC-seq peaks in bin 10 that lack RelA ChIP-seq peaks might correspond to enhancers that function in RelA-independent manner. To address this question, we individually scrutinized ATAC-seq, RelA ChIP-seq, and RNA-seq tracks at the 57 ATAC-seq peaks in bin 10 that did not coincide with RelA ChIP-seq peaks. Twenty-four of these inducible ATAC-seq genomic regions appeared to reflect read-through transcription from nearby genes rather than discrete ATAC-seq peaks representative of enhancers (data not shown; see Figure B-4B below). The remaining 33 ATAC-seq peaks exhibited elevated RelA ChIP-seq reads, but not to a sufficient extent to meet our peak-calling criteria (data not shown). Thus, although we cannot draw firm conclusions from this analysis, we could not convincingly identify strongly induced ATAC-seq peaks associated with the primary response that do not display evidence of RelA binding.

CRISPR/Cas9 Editing to Mutate NF-κB Motifs and the *Rela* and *Rel* Genes

The unique pattern of enrichment of NF-κB motifs (Figure B-2B) and the extensive binding of RelA to the most strongly and consistently induced ATAC-seq peaks (Figure B-3B) suggest that

NF-κB may be a critical and broad regulator of inducible nucleosome remodeling during the primary response to TLR4 signaling. To test this hypothesis functionally, we first used CRISPR/Cas9 editing with homology-directed repair (CRISPR-HDR) to introduce substitution mutations into two NF-κB motifs underlying a potently induced ATAC-seq peak between the *Tyk2* and *Cdc37* genes (Figure B-S2A). For CRISPR-HDR editing, we used a J2 virus-transformed mouse macrophage cell line that displays a robust transcriptional response to lipid A stimulation (data not shown). Inducible ATAC-seq sensitivity at the *Tyk2/Cdc37* intergenic site was selectively absent in the mutant line (Figure B-3C, tracks 6-10), strongly suggesting that NF-κB binding is critical for the inducible ATAC-seq signal. To confirm that this effect was not due to the inability of reads spanning the mutations to map to the wild-type mouse genome, read numbers at specific local sequences that are not altered by the mutation were examined (Figure B-S2A,B); the results confirmed that the NF-κB binding site mutations eliminated ATAC-seq induction. Importantly, transcriptional induction of the inducible *Tyk2* gene was strongly reduced by the NF-κB motif mutations (Figure B-S2C), demonstrating that *Tyk2* is a target of this putative enhancer. Transcription *of Cdc37*, which is not induced by lipid A in wild-type cells, was unaffected by the mutations (data not shown).

To further examine the role of NF-κB in inducible remodeling, we performed ATAC-seq with lipid A-stimulated macrophages in the presence of the IκB kinase inhibitor, Bay 11. Bay 11 treatment resulted in strong inhibition of the inducible *Tyk2/Cdc27* intergenic ATAC-seq peak (data not shown). Because of Bay11's imperfect specificity, we used CRISPR/Cas9 editing to create mutant versions of the J2-transformed macrophage line in which homozygous deletions were introduced into both the *Rela* and *Rel* (c-Rel) genes ($Rela^{-/-}Rel^{-/-}$ line; Figure B-S3). The combined mutations eliminated the possible impacts of redundancy between the two NF-κB family members. Possible redundancy with the RelB member of the NF-κB family was not addressed (*Relb* mRNA remained low following lipid A stimulation of the mutant line but was induced 10-fold in the wild-type line, consistent with the fact that *Relb* is an NF-κB target [data not shown]). Importantly, the inducible ATAC-seq signal at the *Tyk2/Cdc37* intergenic site was eliminated in

two independent *Rela*$^{-/-}$*Rel*$^{-/-}$ clonal lines (Figure B-3C, tracks 11-15, and data not shown). This result further strengthens the evidence that NF-κB binding is essential for inducible nucleosome remodeling at this primary response ATAC-seq peak. Notably, in the *Rela*$^{-/-}$*Rel*$^{-/-}$ line, *Tyk2* mRNA was not induced by lipid A, but with only a marginal reduction in the constitutive *Cdc37* mRNA (Figure B-S2C).

Broad Impact on Strongly Induced ATAC-seq Signals in *Rela*$^{-/-}$*Rel*$^{-/-}$ Cells

We next performed a broader analysis of the ATAC-seq data obtained with the *Rela*$^{-/-}$*Rel*$^{-/-}$ clones following lipid A stimulation for 0, 15, 30, 60, and 120 min (Figure B-4A). The impact of the mutations was analyzed using 1,001 primary response peaks that were induced >5-fold in the WT macrophage line, after separating those peaks into 10 equal bins on the basis of their statistical significance (p-value) of inducibility using time-courses performed with two independent mutant lines. Figure B-4A displays the impacts of the mutations on ATAC-seq RPKM for peaks in the four bins with the greatest statistical significance. In each bin, approximately 90% of the peaks were strongly reduced (<33% of wild-type). One example of a strongly suppressed peak is shown in Figure B-4B (Peak 1). Most of the peaks in each bin that were suppressed to a lesser extent (orange dots in Figure B-4A) were either, 1. strongly suppressed at early time points, with less suppression at the 120-min time point, suggestive of NF-κB-dependence during the primary response but partial NF-κB-independence during the secondary response (Figure B-4B, Peaks 2 and 3), or 2. a result of read-through transcription that does not appear to reflect an inducible regulatory region (Figure B-4B; see Peak 4, which appears to arise from read-through transcription downstream of the lipid A-inducible *Plk2* gene).

Notably, we cannot completely rule out the possibility of NF-κB-independent ATAC-seq induction associated with the primary response at a few peaks (e.g. Figure B-4B, Peak 5). However, peaks that are difficult to classify are rare and still appear to lack ATAC-seq induction at the earliest time points. Therefore, although NF-κB may be universally important for

nucleosome remodeling during the TLR4 primary response, we can only conclude at this time that it is broadly important.

RNA-seq analysis of the wild-type and two independent mutant *RelA*$^{-/-}$*Rel*$^{+/-}$ lines revealed that the mutations strongly reduced transcription (<33% of wild-type) of 74.5% (70 of 94) of genes induced >5-fold that reached an RPKM threshold of 3 RPKM (Figure B-4C, right). In contrast, only 28.1% (38 of 135) of genes induced to a lesser extent (2-5-fold) exhibited strongly reduced transcription in the mutant lines (Figure B-4C, left). Examples of NF-κB-dependent (*Nfkbiz*) and NF-κB-independent (*Dusp5* and *Ccl7*) genes are shown in Figure B-4D. In our prior efforts to classify TLR4-induced genes on the basis of their regulatory requirements (Tong et al., 2016), *Dusp5* was classified as a serum response factor (SRF) target gene and *Ccl7* was among an unusual class of genes that exhibited TRIF-dependence but MyD88- and IRF3-indepencence. Importantly, a manual inspection of IGV tracks surrounding RelA/c-Rel-independent genes failed to identify strongly induced ATAC-seq peaks (data not shown), suggesting that the induction of these genes does not require inducible nucleosome remodeling.

Due to the absence of a definitive method for identifying bona fide nucleosome remodeling events and the large size of the genome, we remain unable to conclusively determine whether or not inducible nucleosome remodeling during the primary TLR4 response can occur at a few locations in the absence of NF-κB. Nevertheless, our results provide strong support for a model in which NF-κB is broadly required for inducible nucleosome remodeling during the primary response.

Critical Role for IRF3 in Inducible Remodeling at a Small Number of ATAC-seq Peaks

The above results do not address whether NF-κB acts alone to promote remodeling or acts in concert with other DNA sequence-specific transcription factors. As shown above, we found enrichment of motifs for several other transcription in bins containing the most strongly induced ATAC-seq peaks, consistent with prior evidence of collaboration between factors (Comoglio et al., 2019). To functionally examine whether NF-κB collaborates with other inducible factors, we

focused first on IRF3, a key transcription factor induced by the TRIF signaling pathway during the primary response to TLR4 signaling.

We previously showed that IRF3 plays a dominant role in the activation of only nine of 132 primary response genes potently induced by TLR4 signaling (Tong et al., 2016). Remarkably, IRF3 was required for nucleosome remodeling only at one promoter, the *Ccl5* promoter, which contains adjacent IRF3 and NF-κB binding sites (Tong et al., 2016). RelA binding was strongly reduced in *Irf3*$^{-/-}$ BMDMs, and RelA bound with slower kinetics to the *Ccl5* promoter than to other primary response promoters, with RelA binding and *Ccl5* induction coinciding with the kinetics of IRF3 activation (Tong et al., 2016). These results suggested that IRF3 is critical for nucleosome remodeling at the *Ccl5* promoter. However, NF-κB's functional role was not examined.

To examine more extensively the extent to which IRF3 contributes to nucleosome remodeling during the TLR4 primary response, and the extent to which it collaborates with NF-κB, we first broadened our analysis of promoters by examining all 40 primary response promoters exhibiting ATAC-seq signals induced >5-fold (see Figure B-S1A). Consistent with our previous results, the *Ccl5* promoter remained unique in its IRF3 requirement among lipid A-inducible ATAC-seq peaks, as demonstrated by both scatter plot and volcano plot analysis comparing wild-type and *Irf3*$^{-/-}$ BMDMs (Figure B-5A, left; stimulations performed in the presence of CHX to focus on the primary response).

Next, to determine how many intergenic primary response ATAC-seq peaks exhibit IRF3 dependence, inducible intergenic ATAC-seq peaks were compared in lipid A/CHX-treated wild-type and *Irf3*$^{-/-}$ BMDMs. In a scatterplot analysis of average RPKM from two independent experiments performed with wild-type and *Irf3*$^{-/-}$ BMDMs (Figure B-5A, right top), only 107 inducible intergenic sites were found to exhibit strong (<33% of WT) dependence on IRF3; 19 of these sites exhibited strong IRF3 binding by ChIP-seq (red dots in Figure B-5A), with another 34 (blue dots) exhibiting weaker binding. A volcano plot analysis (Figure B-5A, right bottom) revealed only 22 statistically significant intergenic, IRF3-dependent, primary response ATAC-seq peaks, 13 of which (59%) strongly bound IRF3, with another 6 exhibiting weaker IRF3 binding. Of the

three that exhibited no binding, two coincided with readthrough transcription downstream of *Ccl4* and *Ccl5* (data not shown).

Notably, IRF motifs were not among the motifs found to be enriched during the primary response in our initial ATAC-seq analysis (see Figures B-2C). However, motif analysis performed with the 107 IRF3-dependent ATAC-seq peaks revealed a clear enrichment of IRF motifs (Figure B-5B). This finding suggests that IRF motifs were not enriched in our initial analysis because IRF3 is important at only a small number and percentage of inducible ATAC-seq regions.

Transcription factors that bind bZIP motifs are difficult to functionally examine due to the large number of proteins capable of binding bZIP recognition motifs, and extensive redundancy between these proteins. However, many bZIP proteins are induced by mitogen activated protein kinases (MAPKs). We, therefore, performed ATAC-seq experiments in the presence of MAPK inhibitors (using a combination of ERK and p38 MAPK inhibitors) to gain preliminary insight into the frequency with which strongly induced primary response ATAC-seq peaks require MAPK signaling. This analysis identified 6 promoters with strongly induced ATAC-seq peaks at which ATAC-seq induction was strongly dependent on MAPK signaling (Figure B-S4). Among 757 primary response intergenic sites that exhibited strongly induced ATAC-seq peaks in these experiments, 166 (21.9%) exhibited strongly reduced ATAC-seq induction in the presence of the MAPK inhibitors (Figure B-S4). Thus, NF-κB appears to collaborate more frequently with MAPK-induced proteins than with IRF3 to promote nucleosome remodeling. However, like IRF3, MAPK-induced proteins are important for only a moderate fraction of strongly induced ATAC-seq events. Additional experiments will be required to determine whether NF-κB acts alone or collaborates with other factors at genomic regions exhibiting strong NF-κB-dependent ATAC-seq induction at which neither IRF3 nor MAPK signaling is required.

Analysis of the *Ccl5* Locus

To further interrogate the possible collaboration between NF-κB and IRF3, we focused on the *Ccl5* locus, which contains two of the high-confidence ATAC-seq peaks exhibiting strong IRF3 dependence; one at the *Ccl5* promoter and an intergenic site located 10 kb upstream of the *Ccl5*

TSS (Figure B-6, compare tracks 3 and 4, IRF3-dependent peaks highlighted with red and blue shading). Both peaks are resistant to CHX (Figure B-6A, compare track 2 and 3), consistent with their classification as primary response peaks, and both are absent in ATAC-seq experiments from *Irf3*$^{-/-}$ BMDMs (Figure B-6A, track 4). Notably, additional lipid A-induced peaks reside upstream of the *Ccl5* promoter, but without IRF3-dependence (Figure B-6A, tracks 1-4).

Both IRF3 and RelA can bind the two IRF3-dependent peak regions following lipid A stimulation (Figure B-6A, tracks 18-21; note that IRF3 binding is also detected at other peaks that do not exhibit IRF3-dependence). Furthermore, both peaks have underlying consensus binding motifs for both IRF3 and NF-κB motifs, with the IRF3 and NF-κB motifs separated by 55 bp at the promoter and by 28 bp at the intergenic region (Figure B-6B and Tong et al., 2016)

Importantly, all of the inducible ATAC-seq peaks, including the two IRF3-dependent peaks, were absent in experiments performed with the lipid A-stimulated *Rela*$^{-/-}$*Rel*$^{-/-}$ mutant lines (Figure B-6A, tracks 5-12). *Ccl5* mRNA induction was also strongly reduced in the *Rela*$^{-/-}$*Rel*$^{-/-}$ lines (Figure B-6A, tracks 16-18) in addition to *Irf3*$^{-/-}$ BMDMs (tracks 13-15). Thus, IRF3 and NF-κB are both required for the inducible ATAC-seq peaks at the two IRF3-dependent sites, revealing that the two factors collaborate to promote remodeling.

It is important to note that the *Ccl5* promoter and distal regions represent the clearest examples of IRF3-dependent ATAC-seq peaks. A small subset of the remaining primary response ATAC-seq peaks that exhibit strong IRF3 dependence are in the vicinity of other inducible primary response genes, and a few of these genes exhibit modest IRF3 dependence (data not shown). However, the induction magnitudes for these genes are small (generally 2-8-fold) in comparison to the induction magnitude of *Ccl5* (>800-fold), many do not exhibit IRF3-dependent induction, and the degree of IRF3-dependence of those genes that exhibit dependence is generally limited (data not shown).

Extended Analysis of the *Ccl5* Locus

To further examine the *Ccl5* intergenic site, we first performed circularized chromosome conformation capture (4C) and capture Hi-C (high-throughput conformation capture) experiments. For these experiments, we focused on the DNA region corresponding to the intergenic IRF3-dependent ATAC-seq peak as the target. The results revealed a limited number of interactions between the target and distant sites in unstimulated BMDMs, with a substantial increase in genomic interactions following lipid A stimulation (Figure B-7A). The most extensive, lipid A-inducible interactions were in the vicinity of the flanking *Ccl3* and *Ccl4* genes, which, like *Ccl5*, are strongly induced by lipid A, but with lesser (roughly 2-fold) dependence on IRF3 for their induction, based on RNA-seq analysis of *Irf3*$^{-/-}$ BMDMs (data not shown). In addition, a zoomed in version of the interaction profile revealed lipid A-enhanced interactions in the vicinity of the *Ccl5* promoter and gene, as well as at other nearby genomic regions (Figure B-S5). Thus, the IRF3-dependent intergenic region participates in interactions consistent with a role in inducible transcription.

Finally, we used CRISPR to delete the genomic region corresponding to the IRF3-dependent ATAC-seq peak (Figure B-7B). The *Ccl5* promoter was deleted separately as a control (Figure B-7B). In three independent clones with deletions of the *Ccl5* promoter, *Ccl5* transcription was abolished, as revealed by RT-PCR analysis following lipid A stimulation (Figure B-7C, right). Deletion of the intergenic region also decreased *Ccl5* transcripts 2 hrs post-stimulation, but by only a small magnitude, (Figure B-7C, left). Despite the observed interactions between the intergenic site and the *Ccl3* and *Ccl4* loci, deletion of the intergenic region had no effect on *Ccl3* or *Ccl4* transcripts (Figure B-7C, bottom). Thus, despite the fact that this *Ccl5* intergenic region contains one of only a small number of IRF3-dependent intergenic ATAC-seq peaks, it may have evolved for the purpose of modulating *Ccl5* transcription by only a small magnitude. Alternatively, it may play a much larger role in the regulation of the *Ccl5* gene and possibly the *Ccl3* and *Ccl4* genes in physiological settings we have not examined, or its function may be redundant with the functions of other nearby regulatory regions.

DISCUSSION

TLR-induced nucleosome remodeling was shown almost 25 years ago to occur at select transcriptional control regions in cells of the innate immune system (Weinmann et al., 1999; Agalioti et al., 2000), and was shown a decade ago to occur at thousands of enhancers, (Kaikkonen et al., 2013; Otsuni et al., 2013; Barozzi et al., 2014). However, the transcription factors regulating TLR-induced remodeling have remained elusive. Several lines of evidence presented here strongly suggest that nucleosome remodeling during the primary response to TLR4 signaling exhibits a surprisingly broad and possibly universal dependence on NF-κB, with the selectivity of remodeling resulting in part from collaboration with other inducible factors. Initial evidence of NF-κB's broad role emerged from the unique relationship between NF-κB motifs (in comparison to other known and de novo motifs) and ATAC-seq peaks that exhibit the strongest induction (Figure B-2B). The highly prevalent binding of the NF-κB RelA subunit to these regions added further evidence of a critical role for NF-κB in nucleosome remodeling and was followed by the finding that CRISPR mutagenesis of NF-κB subunit genes and binding motifs greatly reduced strongly induced ATAC-seq peaks.

The above findings were entirely dependent on the study's focus on the most strongly and consistently induced ATAC-seq peaks; a similarly strong association with NF-κB was not observed at the much larger number of ATAC-seq peaks that are induced by small magnitudes. Although it remains difficult to formally conclude that strongly enhanced accessibility at defined genomic sites in vivo coincides with nucleosome remodeling catalyzed by ATP-dependent remodeling complexes, this relationship has long been hypothesized and experimentally supported. Figures B-3A and B-3B provide a clear view of the importance of combining stringent fold-induction and statistical thresholds for the findings of this study. Figure B-3A first shows that the percentage of peaks that exhibit RelA binding within each bin rises dramatically as either the fold-induction or statistical significance increases. However, the percentages of peaks bound to RelA remains relatively modest even in the final fold-induction and p-value bins. As shown in Figure B-3B, the extremely high prevalence of RelA binding becomes apparent only when the

high fold-induction and statistical criteria are combined. In other words, the use of a stringent fold-induction criterion alone is inadequate because of variable statistical reproducibility of peaks that meet this criterion. The use of a high p-value threshold alone is similarly inadequate because many peaks can reach a high threshold with only small magnitudes of ATAC-seq induction (if the weak induction is reproducible).

The dramatic difference in the prevalence of both NF-κB motifs and NF-κB binding between ATAC-seq peaks that are weakly induced versus strongly induced demonstrates that weak ATAC-seq induction is, on average, fundamentally different from strong ATAC-seq induction. The most likely explanation for this difference is that weak ATAC-seq induction often does not represent inducible nucleosome remodeling, but rather inducible changes in transcription factor binding and other small changes in chromatin architecture associated with inducible transcription. Indeed, it is well-known that DNA cleavage by nucleases and transposases can be impacted by transcription factor binding (Buenrostro et al., 2013; Li et al., 2019). The finding that a large fraction of weakly induced ATAC-seq peaks contain called peaks in unstimulated cells (Figure B-1E), accompanied by constitutive binding of a BAF complex (Figure B-1B) adds further support to the hypothesis that a small statistically significant change in an ATAC-seq signal is less likely than a large change to represent nucleosome remodeling. Importantly, although combining stringent fold-induction and statistical criteria when analyzing ATAC-seq data is likely to enrich for genomic regions that undergo nucleosome remodeling events, these combined criteria remain imperfect measures and most likely relate only to the relative probability that an inducible ATAC-seq peak is representative of inducible nucleosome remodeling. Unfortunately, no assay currently exists that can definitively identify inducible nucleosome remodeling events at a genome-wide scale.

The evidence of NF-κB's uniquely broad role in TLR-induced nucleosome remodeling was unexpected because NF-κB frequently activates transcription by binding both promoters and enhancers that possess open, poised chromatin prior to cell stimulation (Gisletti et al., 2010; Heinz et al., 2010; Tong et al., 2016). Nevertheless, a recent study found that NF-κB binding correlated

more closely with TLR-induced nucleosome remodeling than did the binding of several other candidates examined (Comoglia et al., 2019). Additionally, a second recent study found that prolonged activation of NF-κB activation led to induction of a cluster of ATAC-seq peaks that were not induced when NF-κB was transiently activated by TNFα (Cheng et al., 2021). Our study adds direct functional evidence NF-κB's critical importance for nucleosome remodeling, and demonstrates the breadth of its involvement.

Many studies have shown that inducible nucleosome remodeling is orchestrated by the recruitment of BAF (SWI/SNF) complexes, which possess an ATP-dependent subunit that can catalyze enhanced accessibility through the eviction, sliding, or conformational opening of nucleosomes (Hargreaves and Crabtree, 2011; Clapier et al., 2017; Gatchalian et al., 2020). However, the mechanisms by which NF-κB and its collaborators (IRF3 and MAPK-induced factors) contribute to the recruitment and function of BAF complexes remain to be defined. In fact, the mechanism by which nucleosome remodeling is promoted by well-studied pioneer factors remains only partially understood (Frederick et al., 2023), including the mechanisms by which pioneer factors collaborate with remodeling complexes to access nucleosomal DNA.

Early studies demonstrated that NF-κB can bind stably to nucleosome flanks, but not within the nucleosome interior (Lone et al., 2013). However, local NF-κB binding within the nucleosome core was reported more recently (Stormberg et al., 2021). A recent study of the developmental pioneer factor, PU.1, demonstrated that local transcription factor access allows BAF complex recruitment to compacted chromatin in vitro, with the BAF complex in turn stabilizing PU.1 binding (Frederick et al., 2023). The mechanistic relationship between nucleosome remodeling directed by pioneer factors during development and stimulus-directed nucleosome remodeling remains to be determined, and it therefore may be premature to refer to NF-κB as a pioneer factor.

Additional mechanistic uncertainty arises from the finding that TLR-induced remodeling requires, in at least some instances, a partnership between NF-κB and additional factors (i.e. IRF3 or MAPK-induced factors like AP-1). Among the unanswered questions is why NF-κB

appears to be so broadly required for remodeling during the primary response, with its partners acting at a limited subset of genomic regions? Perhaps NF-κB's broad importance reflects an ability to transiently access its binding sites in nucleosomal DNA and promote binding of its partners, with domains of both transcription factors subsequently needed for the productive recruitment of the BAF complexes. Another unanswered question is whether TLR-induced nucleosome remodeling always requires another DNA-binding protein to partner with NF-κB.

Another unexpected finding was the exceedingly rare importance of IRF3 for nucleosome remodeling during the TLR4 primary response. We previously found an important role for IRF3 in nucleosome remodeling at only one TLR4-induced promoter, the *Ccl5* promoter (Tong et al., 2016). However, a much larger number of IRF3-dependent intergenic remodeling events was expected. Instead, only a small number were identified, most of which are of questionable significance, with the most compelling example located upstream of *Ccl5*. The dependence of *Ccl5* transcription on nucleosome remodeling events dependent on both IRF3 and NF-κB may provide initial insight into the biological relevance of inducible nucleosome remodeling. Our prior nascent RNA-seq experiments revealed that *Ccl5* is among the most potently induced genes in response to TLR4 signaling. The presence of nucleosomes in resting macrophages is likely to constrain aberrant transcription and allow transcription to be activated to a high level only in response to a select subset of stimuli capable of activating both IRF3 and NF-κB.

Finally, our findings suggest that NF-κB plays dual roles in promoting nucleosome remodeling and also as a conventional activator of transcription at control regions already assembled into open chromatin. Whether these two fundamentally distinct functions of NF-κB rely on the same protein domains and whether they are regulated by similar or distinct mechanisms (e.g. post-translational modifications) remains to be determined. Any differences in the regulation of the two functions would open new opportunities for the selective modulation of inflammatory gene expression.

LIMITATIONS OF THE STUDY

A primary limitation of the study is that the ATAC-seq method, even when analyzed with the high-stringency approaches described here, cannot conclusively identify genomic regions subject to inducible remodeling. A second limitation is that, despite the frequent recruitment of BAF complexes to genomic regions displaying strongly induced ATAC-seq peaks, and despite the physical evidence of increased accessibility provided by the ATAC-seq data, we cannot formally conclude that strong ATAC-seq induction is due to nucleosome remodeling events like those that have been extensively studied in vitro.

ACKNOWLEDGMENTS

We thank A. Hoffmann for helpful discussions and critical reading of the manuscript. This work was supported by R01GM086372, R01AI073868, and P50AR063020 (to S.T.S.) T32GM008042 (to B.J.T.), T32AI007323 and F31AI157267 (to A.E.D.). S.T.S. is a Senior Scientific Officer with the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

Conceptualization, S.T.S, A.C.F., and B.J.T.; methodology, A.C.F., B.J.T. and F.S.; formal analysis, A.C.F., B.J.T., and F.S.; investigation, A.C.F., B.J.T., X.L., A.E.D., F.S., and L.C.; writing-original draft, A.C.F. and S.T.S. ; visualization, A.C.F., B.J.T., and F.S.; supervision, S.T.S., P.O.S., and M.F.C., funding acquisition, S.T.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

C57BL/6 was purchased from Jackson Laboratory (Bar Harbor, ME), and *Irf3*$^{-/-}$ mice were a gift from Genhong Cheng (UCLA). Catalog numbers and specific sources of the mice are provided in the Key Resources Table. Mice were bred and maintained on the C57BL/6 background in the Smale laboratory vivarium. The mice were monitored by the UCLA Division of Laboratory Animal Medicine. Experiments were performed under the written approval of the UCLA Chancellor's Animal Research Committee (ARC) in accordance with all federal, state, and local guidelines.

Primary Cell Culture

BMDMs were prepared from 8 to 10-week-old C57BL/6 and *Irf3*$^{-/-}$ male mice as described previously (Ramirez-Carrozzi et al., 2009; Bhatt et al., 2012).

Cell Line

J2 virus-immortalized macrophages were also from C57BL/6 mouse and named as B16#5 by former colleague Xin Liu. B16#5 cells were cultured in the same media for BMDMs and treated accordingly with the same dose of lipid A and CHX for downstream assays.

METHOD DETAILS

Stimulation

The BMDMs or macrophage cells were activated with 100 ng/ ml lipid A (Sigma, L6895) on day 6 for the indicated time course. When indicated, cells were preincubated with 10 ug/ mL CHX (Sigma, 239765) for 15 minutes or with 10 uM Bay11 (Sigma, B5556) for one hour.

ATAC-seq

ATAC-seq libraries were prepared using the Nextera Tn5 Transposase kit (Illumina) as described (Buenrostro et al., 2015) with slight modifications. In brief, approximately 50,000 cells were washed with PBS twice and lysed with a cold lysis buffer (10mM Tris pH 7.4, 10 mM NaCl, 3 mM MgCl2, and 0.1% Igepal) followed by centrifugation at 500 g to obtain nuclei. Cell nuclei were suspended in a 50 ul reaction buffer containing 25 ul 2x Tagment DNA buffer, 2.5 ul Tagment

DNA enzyme (Tn5), and 22.5 ul nuclease free water. Transposase reaction was carried out at 37°C for 30 minutes on the thermomixer with 700 rpm intermittent mixing and cleaned up immediately with Qiagen MinElute PCR purification kit to generate 20 ul transposed DNA. Resulting DNA was PCR amplified for 7 cycles with corresponding barcoding primers and size selected by gel excision 100-1000 bps. Library DNA was purified and quantified by Qbit and Agilent 4200 TapeStation followed by 50 bps single-end sequencing using a HiSeq2000 or a HiSeq3000 sequencing platform.

RNA-seq

Total RNA was extracted from BMDMs or macrophage cells using TRIzol reagent (ThermoFisher #15596018) followed by Qiagen RNAeasy plus mini kit (QIAGEN #74136) according to the manufacturer's instructions. To eliminate genomic DNA contamination, RNA samples were treated with DNase I (QIAGEN #79256) at room temperature for 15 minutes on the RNAeasy column membrane before elution. Total RNA samples were quantified by nanodrop and diluted to 500 ug / 50 ul. Libraries were prepared with Illumina TruSeq RNA Library Preparation Kit 2.0 (Illumina #RS-122-2001) and sequenced on Illumina Hiseq 2000 or 3000 for 50 base-pair single-end reads.

ChIP-seq

ChIP-seq for Rel-A, IRF-3 (Santa Cruz # sc-9082 X) and Brg1 (Cell Signaling # 9284S) was performed as described (Barish et al., 2010), with optimization for each antibody. ChIP-seq libraries were prepared using the Kapa Hyper Prep Kit (Kapa Biosystems), followed by sequencing (Illumina HiSeq 3000). ChIP-seq was performed as previously described (Barish et al., 2010).

CRISPR-Cas9 Ribonucleoprotein (cRNP) Genome Editing and Deletion

A. Guide RNA Design

Two guide RNAs were designed for each deletion or edit site by using CRISPOR online tool (Concordet and Haeussler, 2018). The gRNAs were ranked from highest to lowest specificity

score (Hsu et al., 2013) and gRNAs satisfied our need with the highest specificity score were selected. The selected gRNAs were synthesized and purchased from SYNTHEGO.

B. HDR Template Design and Preparation

The HDR template was designed as sense single-stranded DNA containing 300 nucleotides flanking the replaced sequence and purchased from IDT in the amount of 9 µg. The dry synthesized DNA oligo was dissolved in 4 µL resuspension buffer R provided with Neon™ Transfection System 10 µL kit (Thermo Fisher Scientific, Catalog #MPK1025) right before electroporation.

C. Assemble of cRNP complex

Cas9 2NLS nuclease (*S. Pyrogenes*) was purchased from SYNTHEGO as 20 uM stock and the dry synthetic gRNAs (sgRNAs) were prepared into 30 µM stock. We assembled cRNP complexes by mixing Cas9 10 pmol and sgRNAs 90 pmol in 3.5 µL of resuspension buffer R, then incubate the mixture for 10 minutes at room temperature. For NF-κB motif editing, Cas9 and sgRNAs were added in 3.5 µL of resuspension buffer R containing HDR templates.

D. Prepare Cells

B16#5 cells were subculture one day before electroporation in an appropriate number so that they were about 60-70% confluent on the day of transfection. Before electroporation, the cells were scraped off from the 10 cm plate and counted to obtain 1.5×10^6 cells enough for 10 transfections. The cells were washed with 1x PBS and resuspended in 50 µL of resuspension buffer R.

E. Electroporation

To prepare for a cell-RPN solution, 5 µL of cell suspension was added into the 7 µL RNP mix (total 12 µL). Electroporation was performed using the Neon Transfection System at pulse code (10 ms x 4 pulses) using 10 µL Neon tips at 1,900V. Immediately following electroporation, cells were pipetted into a 6-well plate containing prewarmed culture media.

F. Single Cell Colony Expansion

Fresh cell culture media was added to replace the media containing dead cells 24 hours after electroporation. Forty-eight hours after electroporation, cells were scraped off from the plate, counted, and seeded 2 cells/well in the 96-well plates. Wells containing single-cell colonies were selected under a microscope and replaced with fresh media every four days until wells turned yellow. Cells were then detached from the wells with trypsin and collected for freezing storage and DNA extraction.

G. Screening and Sequencing

Enhancer and promoter deletion clones: DNA was extracted from the single-cell colonies and PCR amplified regions containing deleted sites followed by gel electrophoresis. Clones with successful deletion were selected based on the size of the amplified fragment and confirmed by Sanger sequencing.

NF-κB motif edited clones: The regions containing the edited sites were PCR amplified and successfully edited clones were confirmed by Sanger sequencing directly.

qRT-PCR

RNA was extracted using TRI-reagent (Molecular Research Center, TR118), and treated with DNase I (Qiagen, 79254), and purified using the RNeasy kit (Qiagen, 74106). 1ug of RNA was reverse-transcribed using random hexamers and primers targeting mRNA of *Ccl5*, *Ccl3*, and *Ccl4* designed to amplify products quantified by SYBR-green.

4C Assay

The 4C library was prepared as described (Krijger et al., 2020). In brief, we crosslinked the lipid-A treated BMDMs with 1% Formaldehyde for 10 minutes and quenched them with 2M Tris (0.125M final concentration) for 10 minutes at room temperature. Cells were lysed for 30 minutes in 1 mL cold lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, and 0.2% NP-40 supplemented with protease inhibitor). Nuclei were pelleted and resuspended in 500 µL of 1st restriction enzyme

buffer with 0.1% SDS and incubated for 1 hour at 37°C at 1200rpm. SDS was then quenched by adding Triton-X 100 (final concentration 0.8%) for another 1 hour. 400 units of 1st restriction enzyme (DpnII, NEB R0543M) was added to samples, incubated at 37 °C for 6 hours following by adding additional 400 units for overnight digestion. The 1st restriction enzyme was inactivated by incubating at 65°C for 10 minutes and placed samples on iced immediately. Ligation buffer (66 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 5 mM DTT, 1 mM ATP) was added to dilute samples to a total volume of 7 mL. 2,000 units of T4 DNA ligase was added to each sample and incubated at 16°C overnight to achieve proximal ligation. Reverse crosslinking and RNA digestion were performed and followed by phenol/chloroform DNA extraction. Ligation efficiency was confirmed by running the samples before and after ligation on 1% agarose gel. The ligated DNA was digested overnight in a volume of 500 uL with 200 units of 2nd restriction enzyme (CviQI, NEB R0639L). The digested samples were diluted in the ligation buffer into a total volume of 7 mL. Proximal ligation was performed by adding 4000 units of T4 DNA ligase and incubated overnight at 16°C. The 4C template DNA was extracted by phenol/chloroform and ethanol precipitation. The extracted DNA was further cleaned up using the QIAquick PCR Cleanup kit (QIAGEN, 28104).

To generate the library for subsequent sequencing, reading and non-reading primers were designed for each viewpoint. Choice-Taq™ DNA Polymerase (Denville, CB4050-2) was used to PCR amplify 4C DNA templates containing the viewpoints in four of 25 µL reactions. The PCR products were purified with the illustra™ GFX PCR DNA and Gel Band Purification kit (Sigma, GE28-9034-66) to remove primer dimers. Libraries were sequenced for 50 bp single-end reads on the Illumina Hiseq2000 platform.

Promoter Capture Hi-C Assay

The CHi-C library was prepared as described (Schoenfelder et al., 2018) with slight modifications. Cells were crosslinked in 1% formaldehyde for 10 minutes at room temperature and quenched with 2M Tris-HCl pH 8.0 (final 0.125M) for 10 minutes at room temperature. Cells then were lysed in 300 µL of cold lysis buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.2% Igepal supplemented

with protease inhibitor). The nuclei were digested in 1x NEBuffer 2 containing 100 units of MboI (NEB, R0147L) and incubated for 6 hours at 37 °C following by heat inactivation (65°C for 10 minutes). DNA polymerase I, Large (Klenow) fragment (NEB, M0210L), and biotin-14-dATP (ThermoFisher, 19524016) were used to fill in restriction fragment overhangs and biotinylate DNA ends. 4000 units of T4 DNA ligase then added for proximate ligation and incubated at 16°C overnight. After reverse crosslinking, DNA was extracted with phenol/chloroform and ethanol precipitation. Sonication was then performed to make the biotinylated DNA size suitable for high-throughput sequencing (300-500 bp). DNA was dissolved in 130 µL 10 mM Tris-HCl pH 8.0 and transferred to microTUBE (Covaris, 520045) for sonication (Peak Incidence Power 50W, Duty Factor 10%, 200 Cycle per Burst, 60 seconds). Sonicated DNA was size selected for fragments in the range of 300-500 bp by AMPure XP purification beads (VWR, 75803-122). The fragments marked with biotin were pulled down with Dynabeads™ MyOne™ Streptavidin T1 beads (ThermoFisher, 65601). End repair, A-tailing, and paired-end Agilent adaptor ligation (SureSelect XT Library Prep Kit ILM) were performed and followed by on-bead PCR amplification before hybridizing to the customized probe library (Custom 6-11.9 Mb library for 16 samples). The samples were prepared and hybridized to the capture library according to the protocol provided by Agilent (https://www.agilent.com/cs/library/usermanuals/Public/G7530-90000.pdf). The final PCR amplification was carried out for 9 cycles and purified by using AMPure XP beads. The library was paired-end sequenced for 150 bp on the Illumina Novaseq platform.

FIGURE LEGENDS

Figure B-1. ATAC-Seq Profiling of the TLR4 Response

(A) Maximum fold-induction values (y-axis) for 101,448 called ATAC-seq peaks (x-axis) are shown from datasets obtained with BMDMs stimulated with lipid A for 30, 60 and 120 min, in comparison to unstimulated BMDMs. The 101,448 peaks represent a merger of all peaks that

were reproducibly observed in at least one time point. For this analysis, two (30 and 60 min), three (0 min), or five (120 min) biological replicates were examined. Dotted lines denote the 0.5-fold, two-fold, and five-fold thresholds. Numbers of peaks and percentage of total peaks in each of the fold-ranges are indicated in the table.

(B) The genomic location distribution of all 101,448 ATAC-seq peaks and of the 5,178 peaks induced >5-fold are shown.

(C) ATAC-seq browser tracks are shown for representative promoters with decreased (repressed), unchanged (constitutive), or increased (induced) ATAC-seq signals during the lipid A induction time-course. The tracks at the right show an example of a gene for which the induced ATAC-seq signal spans the entire transcription unit.

(D) ATAC-seq peaks were ranked by maximum fold-change across the lipid A stimulation time-course (in comparison to unstimulated cells) and were then separated into 40 bins with 2,536 peaks/bin. The percentage of peaks at eight different genomic locations (y-axis) in each bin (x-axis) are displayed. Induction ranges are shown (top).

(E) ATAC-seq peaks were separated into 40 bins as in panel D. The percentage of peaks (y-axis) with a statistically called peak (blue) or lacking a peak (orange) in unstimulated cells are shown for each bin.

(F) ATAC-seq tracks near a representative inducible gene (*Ccl9*) are shown for unstimulated and stimulated (120 min) cells to emphasize the distinction between strongly and weakly induced peaks.

(G) ATAC-seq peaks were separated into 40 bins as in panel D. The percentage of ATAC-seq peaks (y-axis) in each coinciding with statistically called Brg1 ChIP-seq peaks are shown from ChIP-seq analysis performed in both unstimulated and stimulated (120 min) BMDMs.

Figure B-2. Transcription Factor Binding Motif Analysis

(A) ATAC-seq peaks were separated into 40 bins as in Figure 1. Transcription factor binding motif enrichment analysis was performed with the peaks in each bin. Motifs were grouped by transcription factor family. Induction ranges are shown (top).

(B) Line graphs show the enrichment of motifs quantified as -log(p-value) (y-axis, blue) and as the percentage of peaks with the motif (y-axis, orange) for four transcription factor families, with values shown for each of the 40 bins from panel A (x-axis).

(C) ATAC-seq was performed with lipid A-stimulated macrophages (0, 30, 60, and 120 min) in the absence and presence of CHX. Peaks induced>5-fold were separated into eight bins on the basis of the sensitivity of peak induction to CHX, with motif enrichment determined for the peaks in each bin. Bin 1 contains peaks with the greatest CHX sensitivity (secondary response) and Bin 8, the greatest CHX resistance (primary response). Given the continuum of sensitivities, precise distinctions cannot be made between the primary and secondary responses, but arbitrary separations are shown.

Figure B-3. Initial Evidence of a Broad Role for NF-κB in Nucleosome Remodeling

(A) ATAC-seq peaks were separated into 40 bins on the basis of either fold-induction (blue) or statistical confidence of induction (-log[p-value], orange). The percentage of ATAC-seq peaks in each bin overlapping a RelA ChIP-seq peak (y-axis) is shown for each bin (x-axis). Numerical percentages are indicated for bin 40.

(B) To further examine the prevalence of RelA binding at high-confidence, inducible ATAC-seq peaks, 3,290 primary response ATAC-seq peaks displaying an average induction >5-fold were separated into 10 bins on the basis of their statistical significance of induction (-log[p-value]). The percentage of peaks in each bin overlapping with a RelA ChIP-seq peak (y-axis) is shown for each bin (x-axis). In bin 40, 82.7% of ATAC-seq peaks overlap with a RelA ChIP-seq peaks. Manual scrutiny of tracks for the 57 peaks that lack an overlapping RelA ChIP-seq peak revealed that 33 show evidence of RelA binding below the statistical threshold used to define a called peak;

the other 24 coincide with broad regions of ATAC-seq sensitivity that appear to represent inducible read-through transcription.

(C) Tracks are shown for a genomic region between the *Tyk2* and *Cdc37* genes containing a strongly induced ATAC-seq peak and an overlapping, strongly induced RelA ChIP-seq peak. ATAC-seq tracks are shown from a wild-type J2 virus-transformed macrophage line (tracks 1-5), a clonal line in which substitution mutations were introduced by CRISPR-HDR into two consensus NF-κB motifs underlying the ATAC-seq peak (tracks 6-10), and a mixture of lines in which the *Rela* and *Rel* genes were simultaneously inactivated by CRISPR deletion (tracks 11-15). RelA ChIP-seq tracks from wild-type BMDMs are also shown (tracks 16-20). Tracks are shown for five time points (noted at left). Black horizontal lines denote the locations of constitutive (left) and inducible (right) called ATAC-seq peaks.

Figure B-4. Broad Impacts on ATAC-seq and mRNA Induction in *Rela*$^{-/-}$*Rel*$^{-/-}$ Cells

(A) ATAC-seq was performed with a wild-type J2-transformed macrophage line and two independent clonal *Rela*$^{-/-}$*Rel*$^{-/-}$ lines generated by CRISPR mutagenesis. 1001 primary response ATAC-seq peaks (defined as primary response with CHX-treated BMDMs) induced >5-fold in the wild-type cell line were separated into ten bins based on statistical significance (-log[p-value]) of maximum induction at the time point at which maximum induction was achieved in two independent experiments. Scatter plots are shown for the four bins with the strongest statistical significance ranges. The scatter plots show ATAC-seq RPKM from wild-type (x-axis) and *Rela*$^{-/-}$ *Rel*$^{-/-}$ (y-axis) cells at the time point that yielded the maximum difference (p-value) between wild-type and mutant. The solid diagonal line represents equal RPKM and the dashed line represents average RPKM in the mutant lines that is 33% of the average wild-type RPKM. Peaks that remain >33% of wild-type in the mutant lines are colored in red. The numbered peaks (1-5) are those whose IGV tracks are shown in panel B.

(B) Browser tracks are shown for a representative constitutive ATAC-seq peak and for five representative inducible ATAC-seq peaks (peaks are numbered in panel A) across the lipid A

stimulation time-course from wild-type and *Rela*$^{-/-}$*Rel*$^{-/-}$ mutant cells (tracks contained merged data from two wild-type experiments or from two independent mutant lines). Peak 1 is from an *Nfkb1* intron. Peak 2 is downstream of the constitutive *Tgds* gene. Peak 3 is upstream of the lipid A-inducible *Ptgs2os2* gene. Peak 4 appears to result from readthrough transcription extending downstream of the lipid A-inducible *Plk2* gene. Peak 5 is from an intron for lipid A-inducible *Dcbld2* gene. Scale ranges are consistent among tracks for each peak, but differ from peak to peak.

(C) The scatter plot show Log2 RPKM from mRNA-seq experiments performed with the wild-type J2 line and two independent *Rela*$^{-/-}$*Rel*$^{-/-}$ mutant lines. The plots show all genes induced 2-5-fold (left) or >5-fold (right) 120-min after stimulation with lipid A (maximum RPKM >3). Values are averages of two experiments with wild-type cells or experiments with two independent mutant lines. The solid diagonal line represents equal RPKM in wild-type and mutant and the dashed line represents 33% RPKM in the mutant samples in comparison to wild-type. Grey dots represent genes whose transcript levels are <33% of wild-type.

(D) Browser tracks are shown for one of many strongly induced genes that exhibit strongly reduced mRNA levels in two independent *Rela*$^{-/-}$*Rel*$^{-/-}$ lines (*Nfkbiz*), along with tracks for two primary response genes (*Dusp5* and *Ccl7*) that remained strongly induced by lipid A in the *Rela*$^{-/-}$*Rel*$^{-/-}$ lines.

Figure B-5. IRF3-Dependence of ATAC-seq Peaks

(A) ATAC-seq was performed with wild-type and *Irf3*$^{-/-}$ BMDMs left unstimulated or stimulated with lipid A for 120 min in the presence of CHX to prevent secondary response proteins from masking IRF3 requirements during the primary response (see Tong et al., 2016). Induced ATAC-seq signals (RPKM) from the wild-type and mutant cells were compared by scatter plot (top) and volcano plot (bottom) for the 40 promoters (left) and 1,494 intergenic (right) ATAC-seq peaks induced >5-fold by lipid A. Peaks overlapping with a strong (peak score >90) IRF3 ChIP-seq peak (red) are distinguished from peaks with weak (peak score <90) IRF3 ChIP-seq peaks (blue) and from peaks lacking IRF3 ChIP-seq peaks (black).

(B) IRF and Rel-homology domain (RHD, NF-κB and NFAT) family member binding motif enrichment is shown for representative primary and secondary response ATAC-seq bins (see Figure 2C) and for IRF3-dependent primary response peaks (right). Because only 107 IRF3-dependent ATAC-seq peaks were identified, wild-type BMDM ATAC-seq peaks were divided into bins with 107 or 108 peaks each, to allow an accurate comparison between enrichment in the wild-type and *Irf3*$^{-/-}$ datasets. Only a subset of the wild-type bins are shown, with bin numbers at the bottom.

Figure B-6. NF-κB/IRF3 Co-Dependence of IRF3-Dependent ATAC-seq Peaks

(A) Browser tracks are shown for the *Ccl5* gene and its upstream intergenic region. ATAC-seq tracks are shown for unstimulated and stimulated wild-type BMDMs (tracks 1 and 2); wild-type and *Irf3*$^{-/-}$ BMDM stimulated in the presence of CHX (tracks 3 and 4); and wild-type and *Rela*$^{-/-}$*Rel*$^{-/-}$ J2-transformed macrophages (tracks 5-12). Chromatin-associated RNA-seq tracks are shown for wild-type BMDMs stimulated in the absence and presence of CHX (tracks 13-14) and for *Irf3*$^{-/-}$ BMDM stimulated in the presence of CHX (track 15). mRNA-seq tracks are shown for wild-type and *Rela*$^{-/-}$*Rel*$^{-/-}$ J2-transformed macrophages (tracks 16-18). RelA and IRF3 ChIP-seq tracks are shown from wild-type BMDMs (tracks 19-22). Lipid A stimulation time points, the presence or absence of CHX, and other information listed above are shown to the left.

(B) IRF3 and NF-κB consensus binding motifs and their spacing near the ATAC-seq summits at the *Ccl5* promoter and IRF3-dependent intergenic peak are shown.

Figure B-7. Analysis of *Ccl5* Intergenic NF-κB/IRF3 Co-Dependent Region

(A) 4C and capture Hi-C results are shown for the region surrounding the *Ccl5* intergenic IRF3-dependent ATAC-seq peak. This genomic site was used as a target bait for the experiments. The black line and gray areas above the track represent the 4C trend line, which reflects the breadth and strength of genomic interactions. The arcs below the tracks represents the most significant

capture Hi-C long-range interactions. Experiments were performed with unstimulated BMDM and BMDM stimulated with lipid A for 120 min. Gene annotations are shown at the top.

(B) The locations of the deletions introduced into the IRF3-dependent ATAC-seq peak regions at the *Ccl5* promoter and *Ccl5* intergenic region are shown (red bars). The regions are displayed relative to ATAC-seq tracks from unstimulated and stimulated BMDMs.

(C) RT-PCR was used to compare the impact of the *Ccl5* promoter deletion on *Ccl5* mRNA levels (right), as well as the impact of the intergenic region deletion on *Ccl5*, *Ccl3*, and *Ccl4* mRNA levels (left). For the intergenic deletion, the results of five independent biological replicates are shown, with the results presented as a percentage in comparison to the mean of the wild-type clone signals. P-values are shown for the *Ccl5* data.

Figure B-S1. Characterization of Lipid A-Induced Promoter and Enhancer ATAC-Seq Peaks

(A) A heatmap shows a 120-min timecourse of ATAC-seq induction for promoter peaks that are induced >5-fold, as well as the nascent transcript RNA-seq induction profiles for the same genes. The heatmap including seven previously defined strongly induced primary response promoters (PRG, Tong et al., 2016) and six previously defined secondary response promoters (SRG, Tong et al., 2016), along with the promoters of twenty other genes that are induced more weakly (Induced) and the promoters of six genes that are not induced by lipid A during the 120-minute timecourse (N.R.)

(B) The graphs show the distribution of distances from ATAC-seq peaks to the nearest gene whose nascent transcripts are induced >10-fold within 120 minutes of lipid A stimulation. Distances are on the x-axis and the percentage of peaks within each of many distance ranges are shown on the y-axis. Separate graphs are shown for ATAC-seq peaks induced >5-fold (top), for peaks induced 2-5-fold (second graph), for constitutive peaks (0.5-2-fold, third graph), and for peaks repressed more than 2-fold (bottom graph). Results are shown for the experimental set of peaks (no shuffle) and for a set in which the peaks were computationally randomized (shuffle). See STAR Methods for analysis details.

(C) Intergenic ATAC-seq peaks induced >5 fold during the timecourse of lipid A stimulation (0, 30, 60, and 120 min) were grouped into six kinetic clusters by K-means analysis (columns 1-4). CHX sensitivity is also shown (column 5). At the right, the average induction kinetics (in RPKM) for peaks within each cluster is shown.

Figure B-S2. Validation of CRISPR Mutant Results for *Tyk2/Cdc37* ATAC-seq Site

(A) The DNA sequence underlying the *Tyk2/Cdc37* intergenic ATAC-seq site is shown (top sequence), along with the mutant sequence introduced by CRISPR-HDR (bottom; altered sequence in purple). Numbers (1-7) indicate the read sequences that were monitored in panel B.

(B) Numbers of raw reads at each of the seven locations indicated in panel A are shown from ATAC-seq experiments performed with the wild-type macrophage line (top), a *Rela*$^{-/-}$*Rel*$^{-/-}$ mutant clone (middle), and an NF-κB binding motif mutant clone (bottom). Notably, ATAC-seq reads with sequences that were not altered by the motif mutation are largely absent in the NF-κB binding motif mutant data (presumably due to the impact of the motif mutations on lipid A-induced ATAC-seq accessibility) and also in the *Rela*$^{-/-}$*Rel*$^{-/-}$ mutant data (due to the impact of loss of NF-κB on lipid A-induced ATAC-seq accessibility). Total ATAC-seq read counts for each clone and each time point are shown at the right.

(C) mRNA levels for *Tyk2* (top left), *Cdc37* (top right), and *Ccl5* (bottom left; *Ccl5* serves as a control) were monitored by RNA-seq in unstimulated and lipid A-stimulated BMDMs (gray), the wild-type macrophage line (blue), and the *Rela*$^{-/-}$*Rel*$^{-/-}$ line (red). RPKM (y-axis) and shown for four time points (x-axis). At the bottom right, normalized transcript levels are shown for *Tyk2* and for *Ccl5* (as a control) in the wild-type macrophage line and in the NF-κB binding motif mutant clones.

Figure B-S3. Validation of the *Rela* and *Rel* Deletions

(A) ATAC-seq tracks for the *Rel* gene confirms the presence of the CRISPR-targeted deletion in two independent *Rela*$^{-/-}$*Rel*$^{-/-}$ clones. No mapped reads are present at the deleted region.

(B) ATAC-seq tracks for the *Rela* gene confirm the presence of the CRISPR-targeted deletion in two independent *Rela*$^{-/-}$*Rel*$^{-/-}$ clones.

Figure B-S4. MAPK Inhibitor Impacts on Lipid A-Induced ATAC-seq Peaks

BMDMs were pre-treated for 60 min at 37ºC with either DMSO or a combination of the ERK inhibitor, PD-0325901, and the p38 inhibitor, BIRB-796 (both dissolved in DMSO and used at a final concentration of 5 μM). Following pre-treatment, samples were stimulated with 100 ng/ml lipid A for 0, 15, 30, 60, or 120 min. 757 intergenic ATAC-seq peaks were identified in these experiments that were induced by >2-fold in at least one of the time points and that were defined as primary response peaks in the experiment described in Figure 2C. 35 primary response promoter peaks induced by >2-fold at one of the time points were also identified. Scatter plots show the effect of the MAPK inhibitors on ATAC-seq induction at the promoter (left) and intergenic (right) peaks. Results are shown from either the 60- or 120-min time points for each peak, using the time point showing the greatest effect.

Figure B-S5. Zoomed-Out and Zoomed-In Views of 4C and Capture Hi-C Interactions with the IRF3-Dependent Intergenic ATAC-seq Peak Region Upstream of the *Ccl5* Gene

(A) A zoomed-out view (repeated from Figure 7A) of 4C and capture Hi-C results is shown for the region surrounding the *Ccl5* intergenic IRF3-dependent ATAC-seq peak. This genomic site was used as a target for the experiments. The black line and gray areas above the track represent the 4C trend line, which reflects the breadth and strength of genomic interactions. The arcs below the tracks represents the most significant capture Hi-C long-range interactions. Experiments were performed with unstimulated BMDM and BMDM stimulated with lipid A for 120 min. Gene annotations are shown at the top.

(B) A zoomed-in view of the above data is shown to reveal interactions with the *Ccl5* promoter region.

Figure B-1. ATAC-Seq Profiling of the TLR4 Response

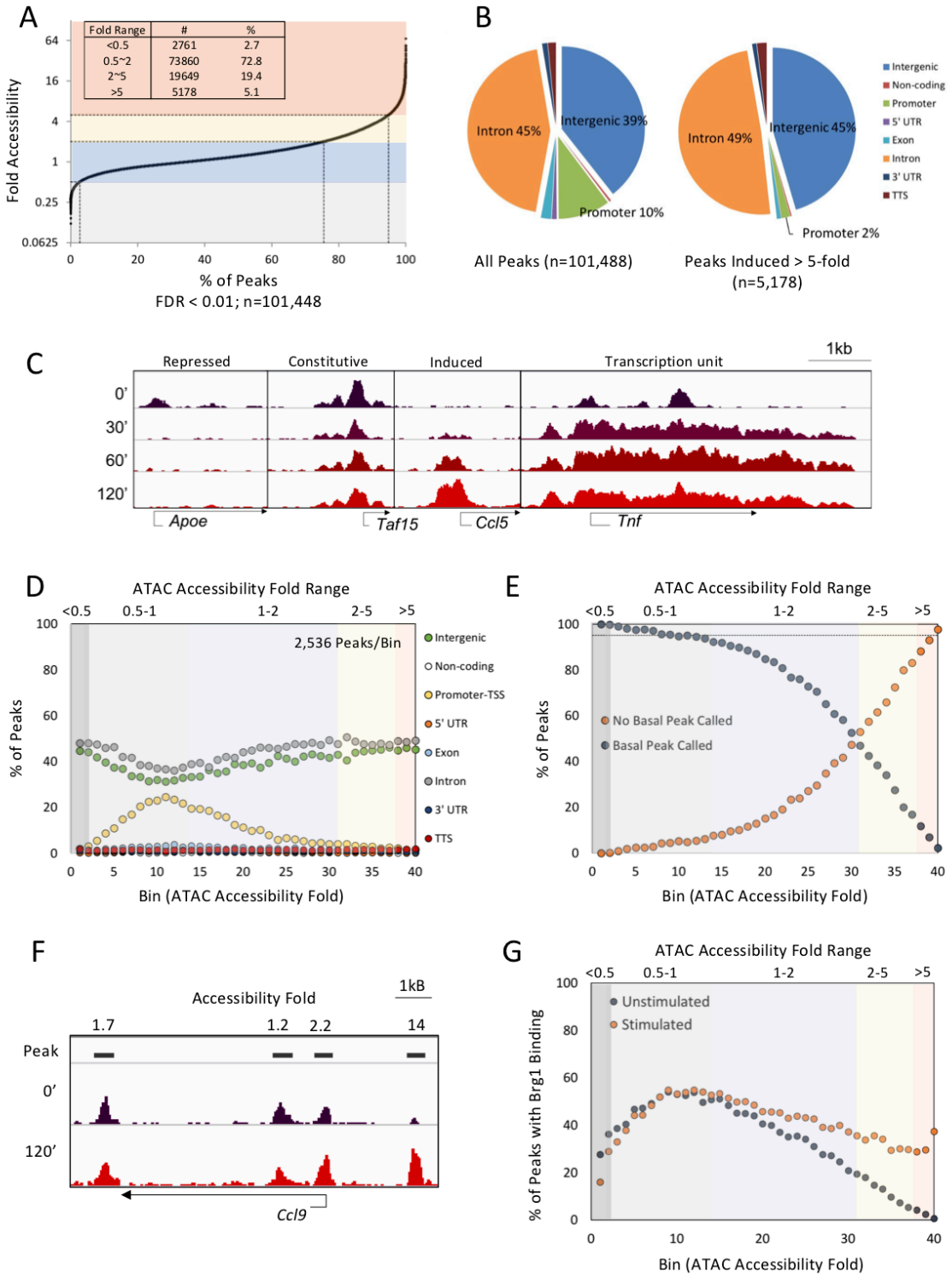


Figure B-2. Transcription Factor Binding Motif Analysis

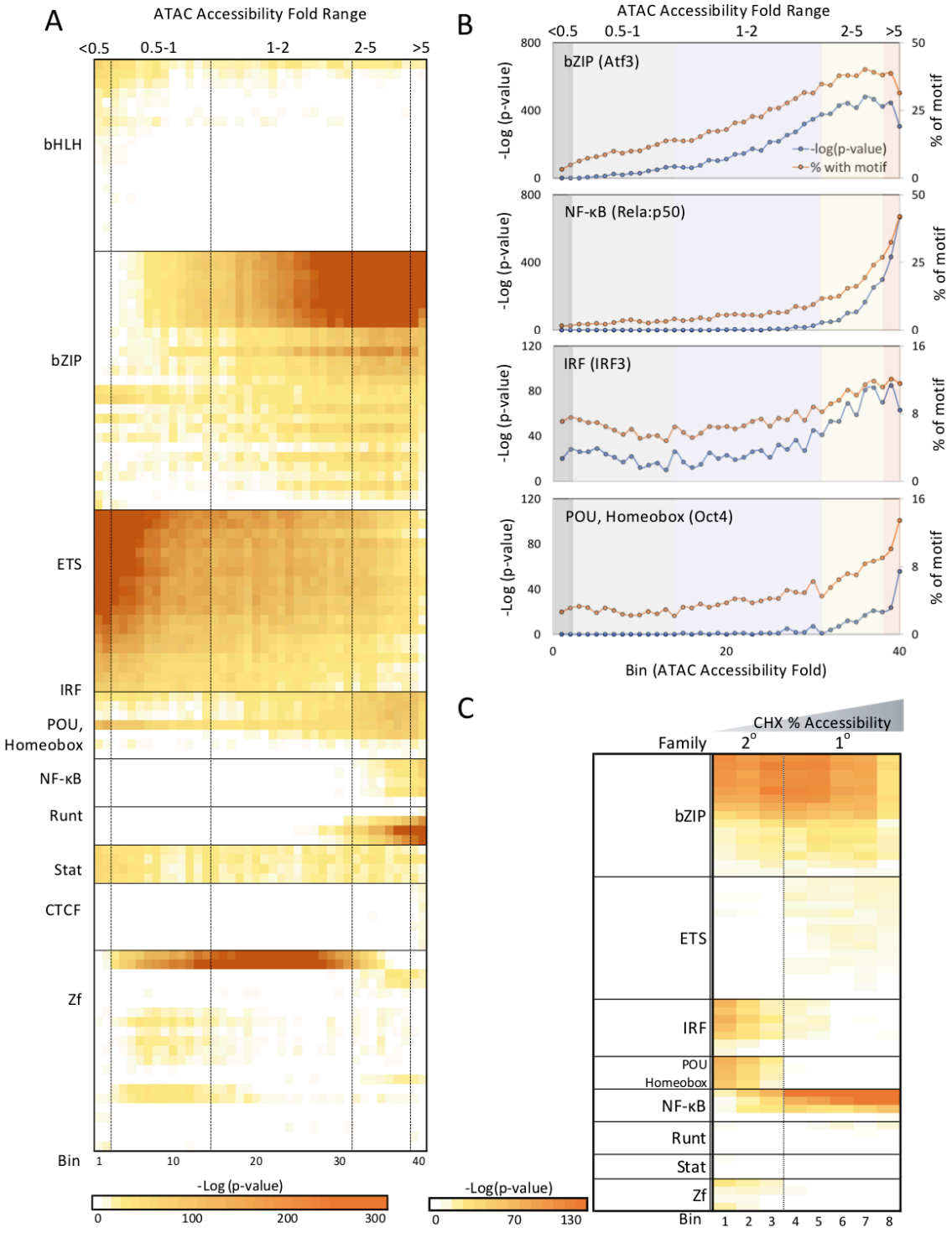


Figure B-3. Initial Evidence of a Broad Role for NF-κB in Nucleosome Remodeling

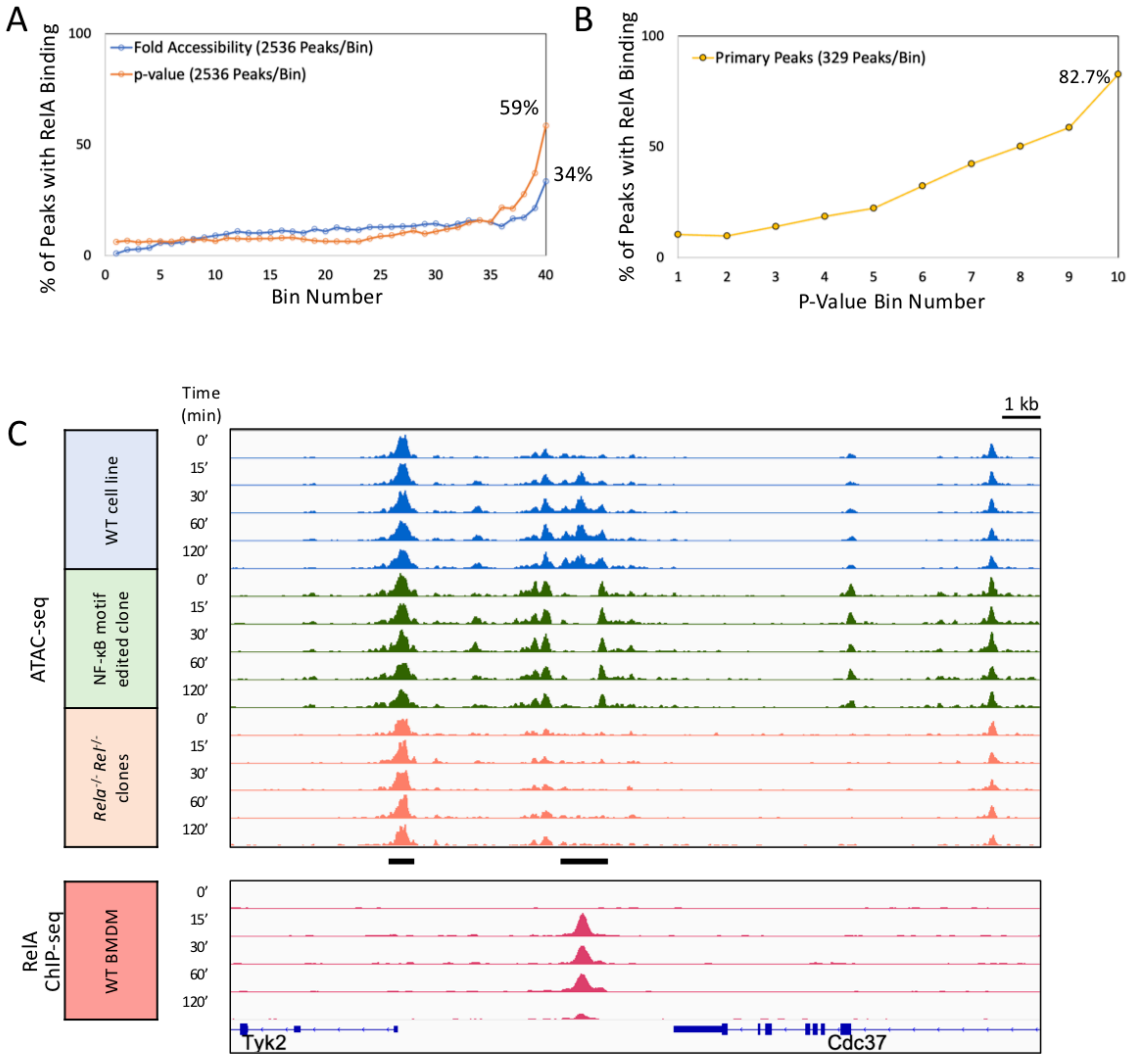


Figure B-4. Broad Impacts on ATAC-seq and mRNA Induction in *Rela*$^{-/-}$*Rel*$^{-/-}$ Cells

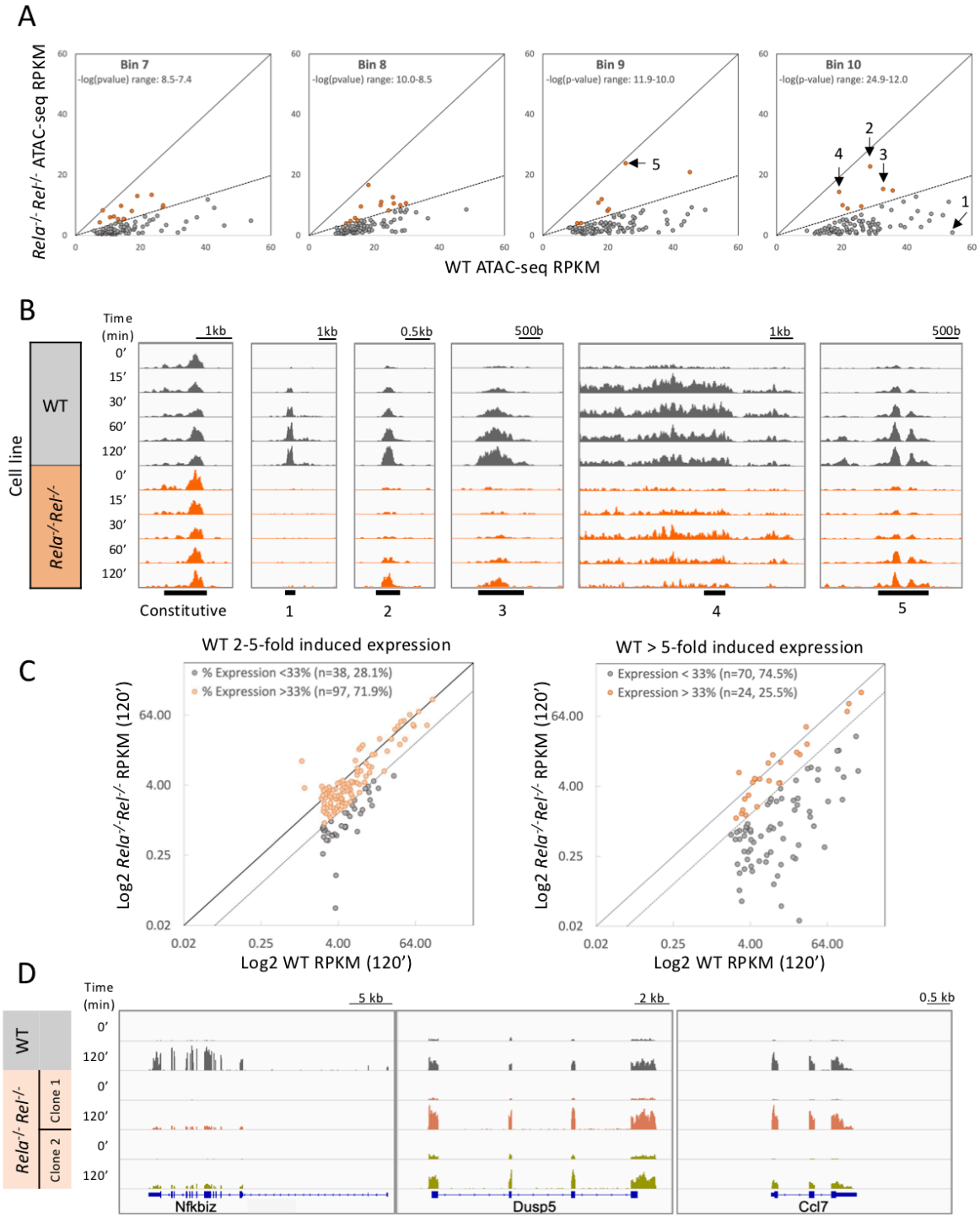
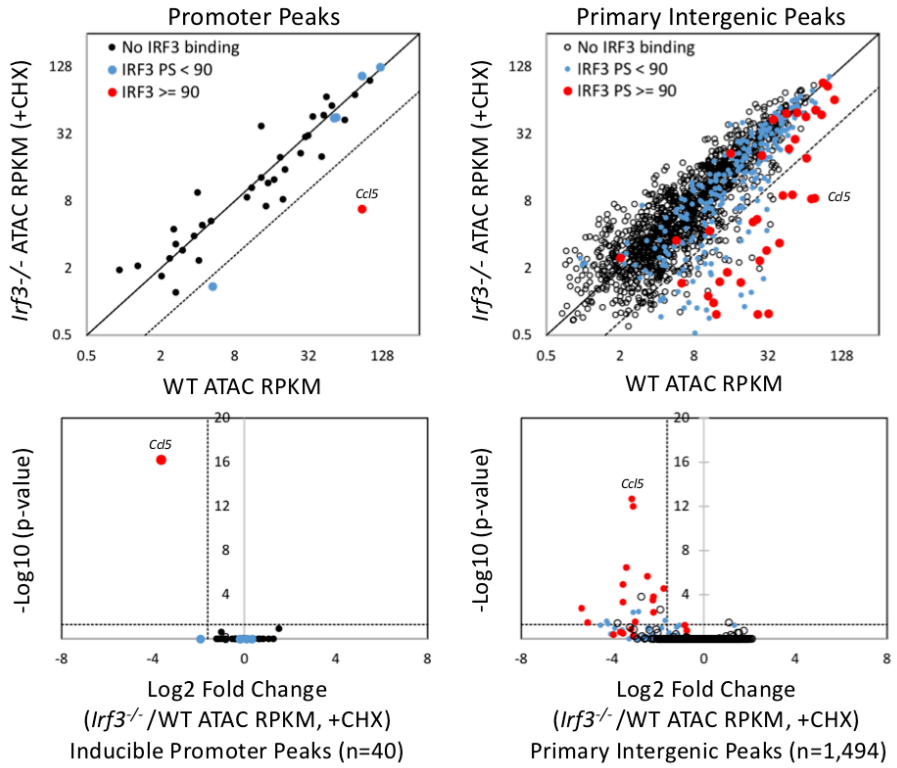


Figure B-5. IRF3-Dependence of ATAC-seq Peaks

A

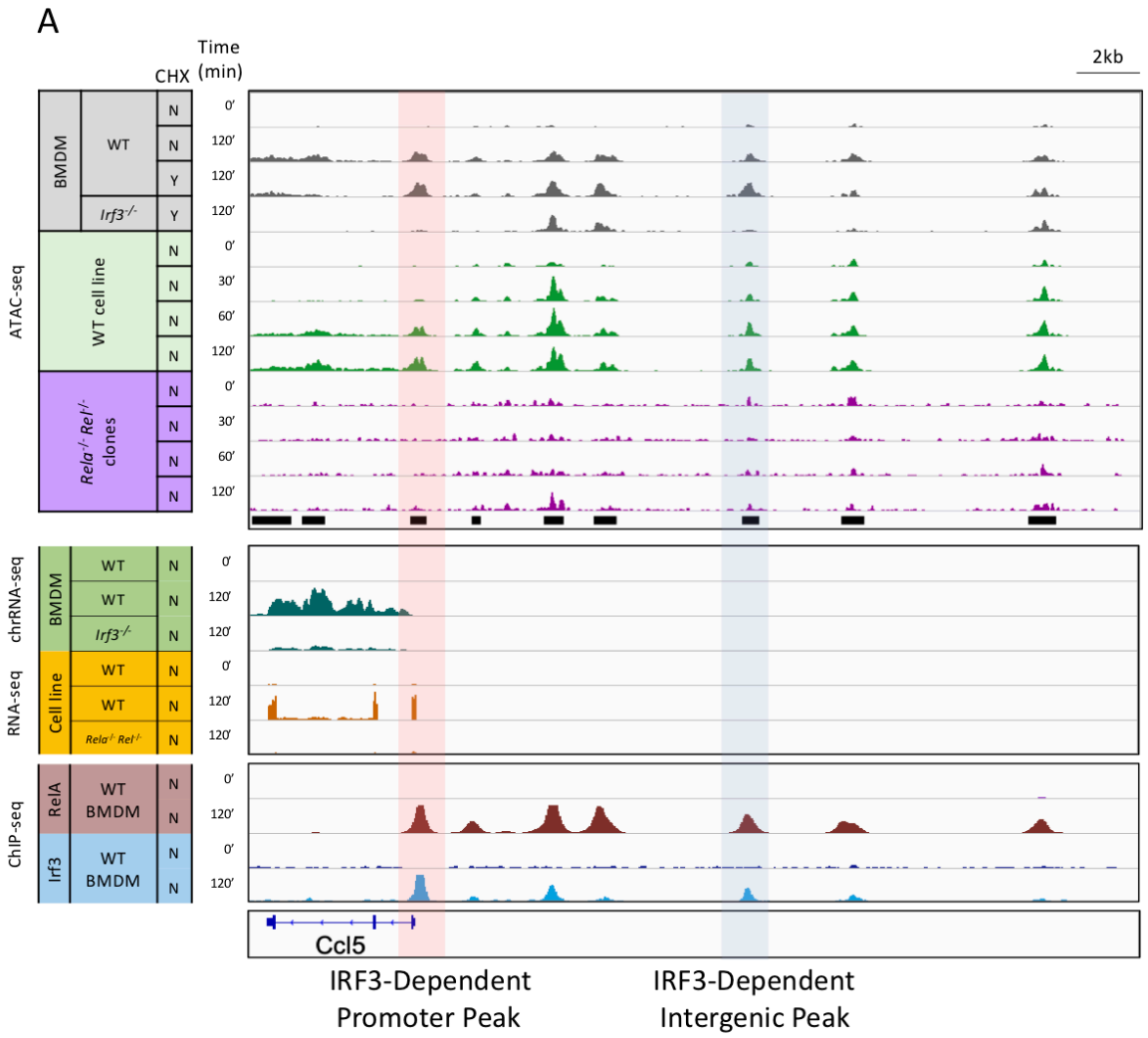


B

Family	Motif Name	Secondary				Primary				Primary IRF3 dependent peaks
IRF	IRF1	17	13	4	5	1	0	0	3	7
	IRF3	15	12	7	3	4	1	1	1	22
	IRF8	15	8	6	4	3	0	0	4	3
	IRF2	13	13	4	2	1	0	1	3	8
	ISRE	8	12	8	4	1	0	0	3	10
	T1ISRE	4	0	0	0	0	0	0	0	0
	IRF4	3	1	0	0	0	0	0	1	2
RHD	NFkB-p65	1	5	11	25	30	39	45	39	12
	NFkB-p65-Rel	0	0	8	21	13	21	29	25	2
	NFAT	0	1	0	1	0	1	0	3	2
	NFkB-p50,p52	0	0	3	5	4	6	6	15	5
	Bin	1	4	7	10	13	16	19	21	

-Log(p-value)
0 25 50

Figure B-6. NF-κB/IRF3 Co-Dependence of IRF3-Dependent ATAC-seq Peaks



B

	IRF Motif		NF-κB Motif
***Ccl5* Promoter Peak**	GAAACTGAAATC	- 55 bp -	GGGAGTTTCC
***Ccl5* Intergenic Peak**	GAAACAGAAACT	- 28 bp -	GGGCTTTGCC

Figure B-7. Analysis of *Ccl5* Intergenic NF-κB/IRF3 Co-Dependent Region

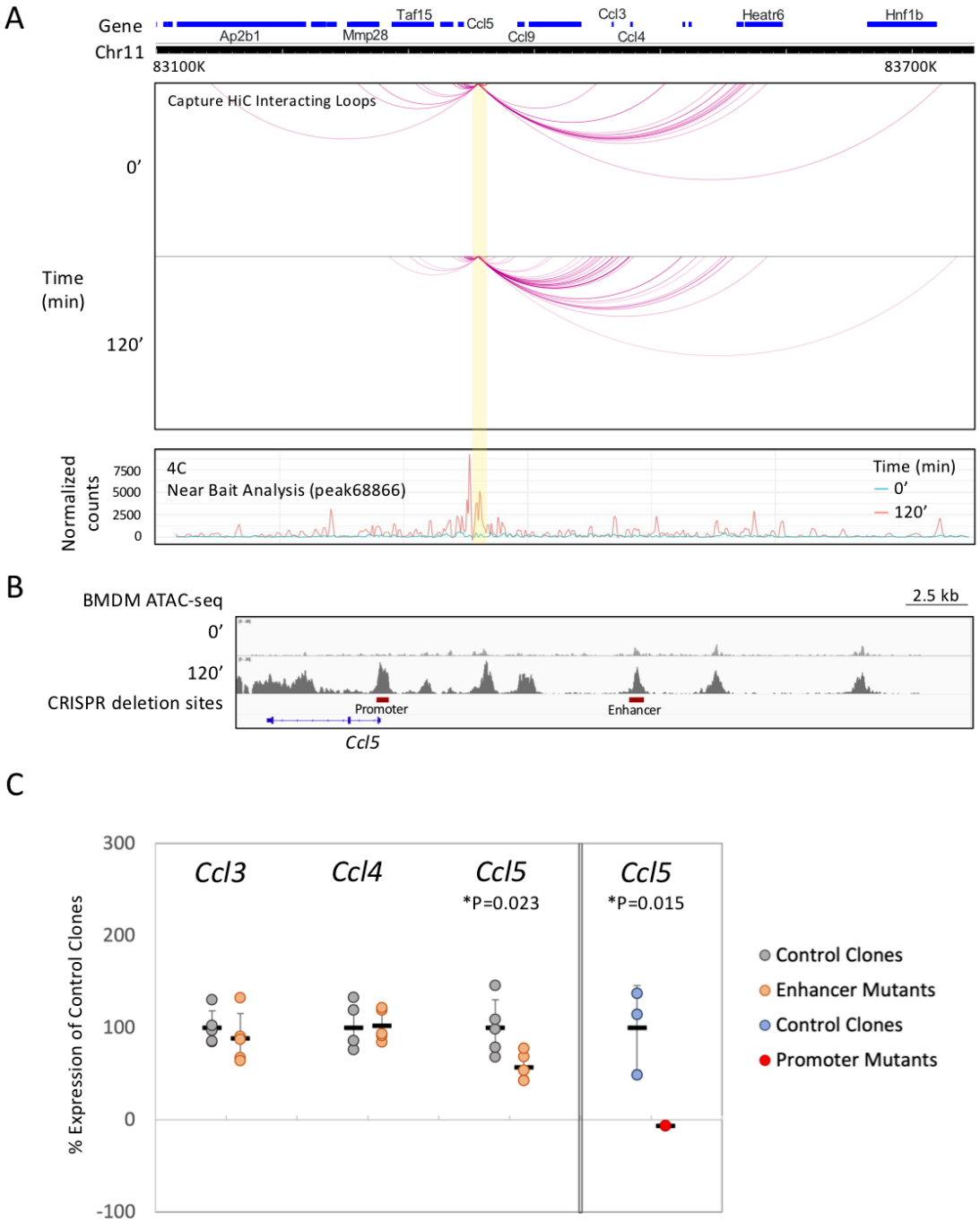


Figure B-S1. Characterization of Lipid A-Induced Promoter and Enhancer ATAC-Seq Peaks

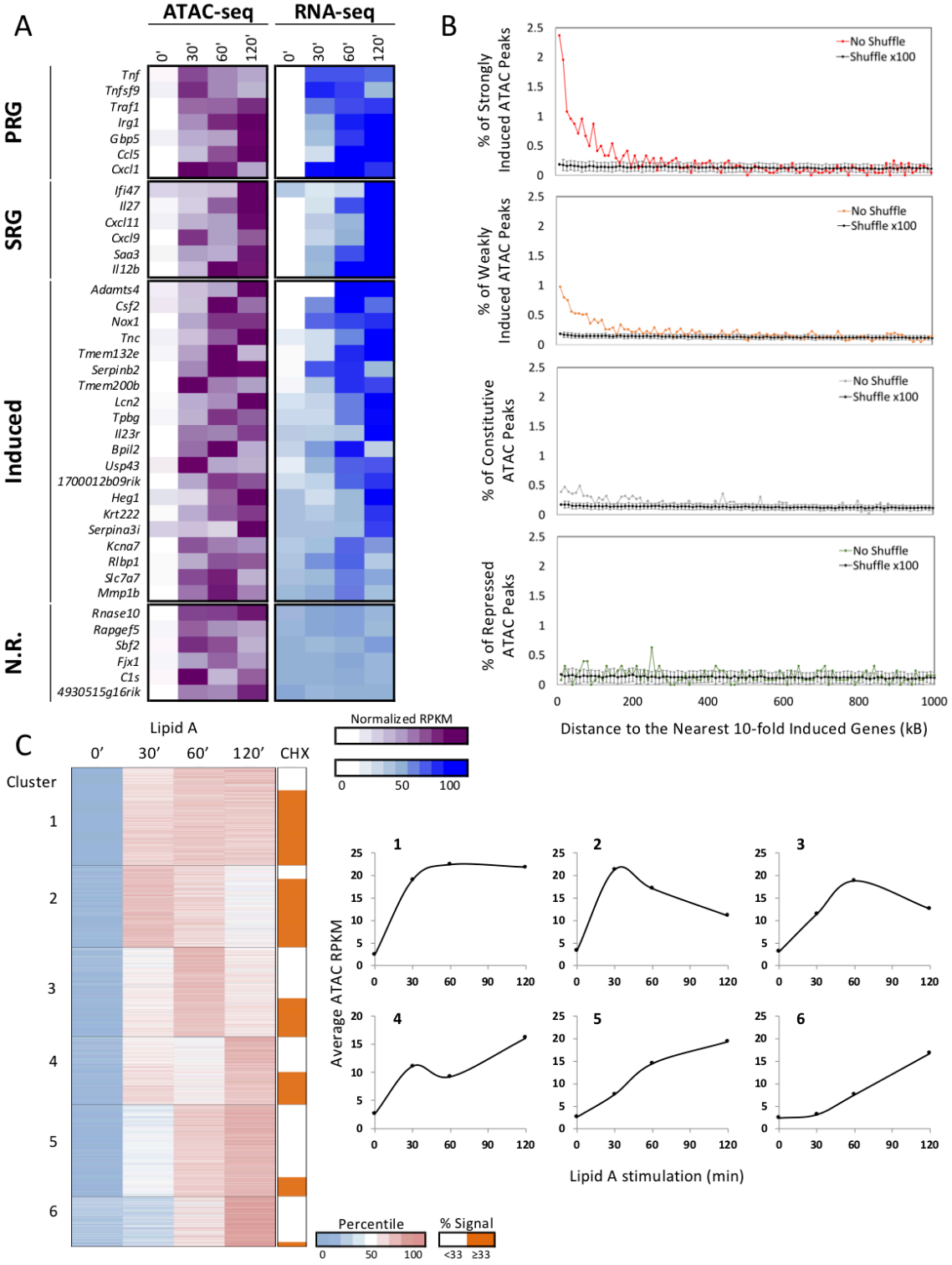
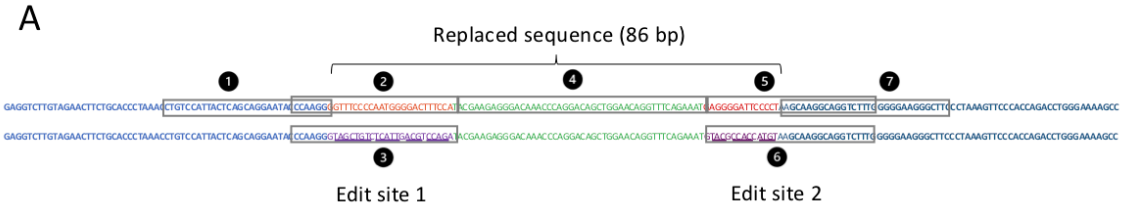


Figure B-S2. Validation of CRISPR Mutant Results for *Tyk2/Cdc37* ATAC-seq Site

A



B

		Seq1	Seq2	Seq3	Seq4	Seq5	Seq6	Seq7	Total reads
	WT_0	5	1	0	0	0	0	2	35,794,772
	WT_15	17	14	0	4	3	0	7	35,738,973
	WT_30	46	12	0	8	4	0	13	42,701,952
	WT_60	37	11	0	9	5	0	11	41,676,140
	WT_120	29	7	0	8	3	0	6	34,348,773
Rela$^{-/-}$ *Rel*$^{-/-}$ clones	DKO_0	2	1	0	1	0	0	0	24,423,149
	DKO_15	3	4	0	1	1	0	2	31,502,927
	DKO_30	3	3	0	1	0	0	1	29,789,571
	DKO_60	3	0	0	1	0	0	1	30,777,515
	DKO_120	5	0	0	0	0	0	1	30,771,543
NF-κB motif mutant	N3_0	0	0	0	0	0	0	0	25,678,270
	N3_15	0	0	0	0	0	0	0	28,271,549
	N3_30	0	0	0	0	0	1	1	21,934,904
	N3_60	0	0	1	1	0	0	0	29,440,559
	N3_120	1	0	0	0	0	0	0	27,280,072

C

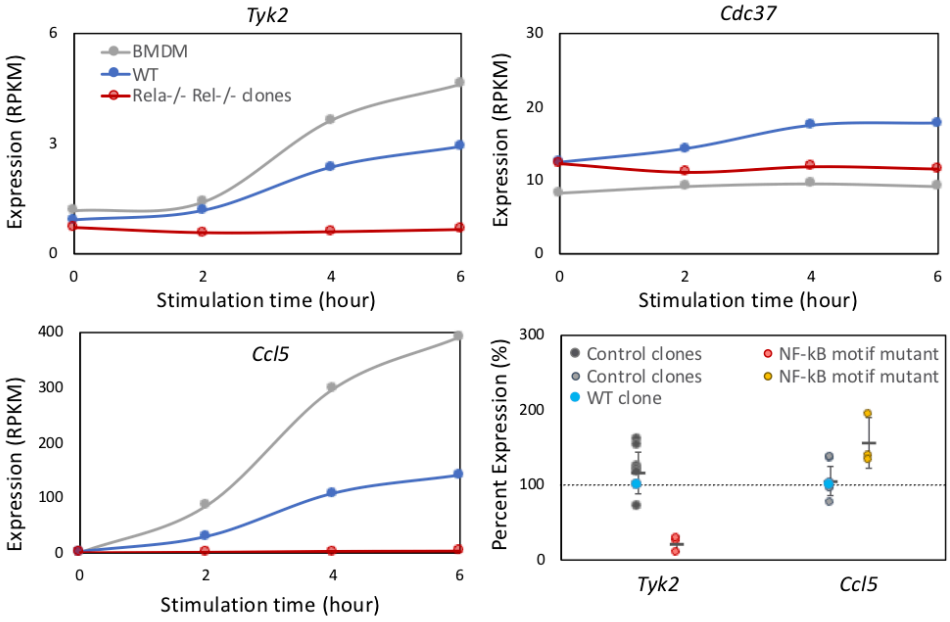


Figure B-S3. Validation of the *Rela* and *Rel* Deletions

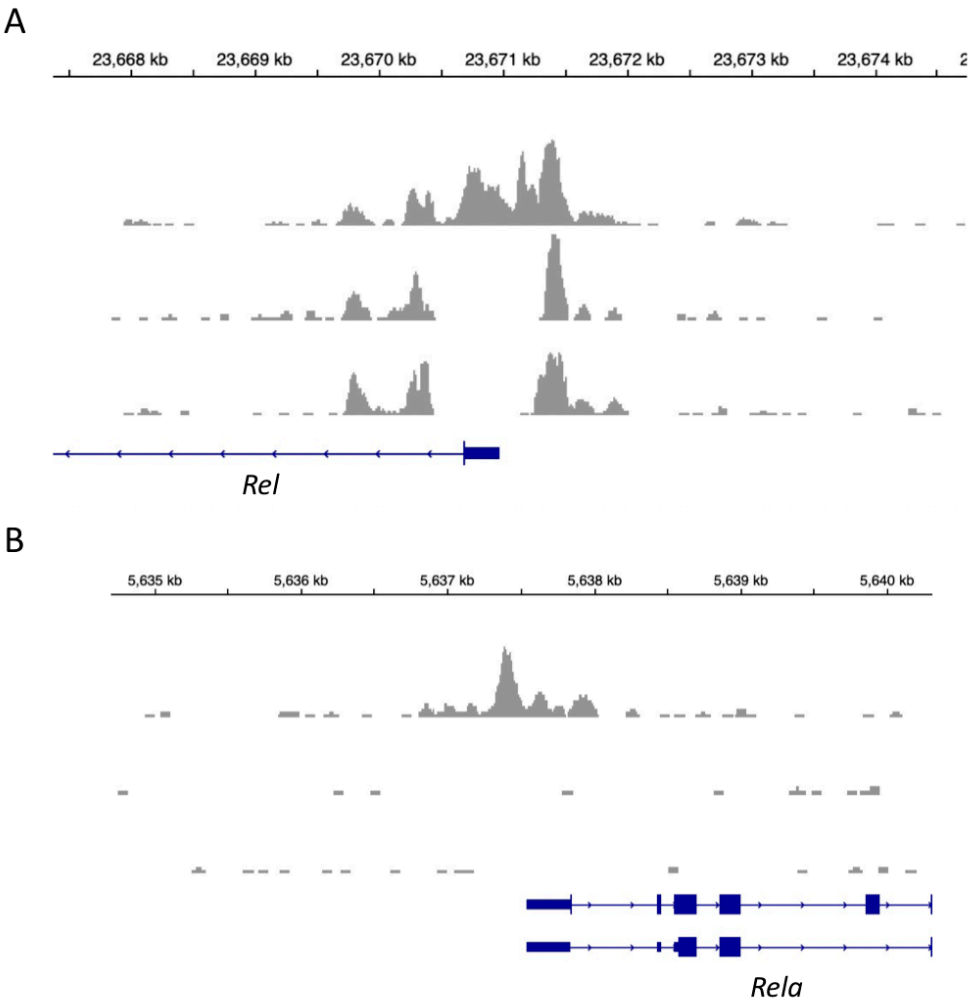


Figure B-S4. MAPK Inhibitor Impacts on Lipid A-Induced ATAC-seq Peaks

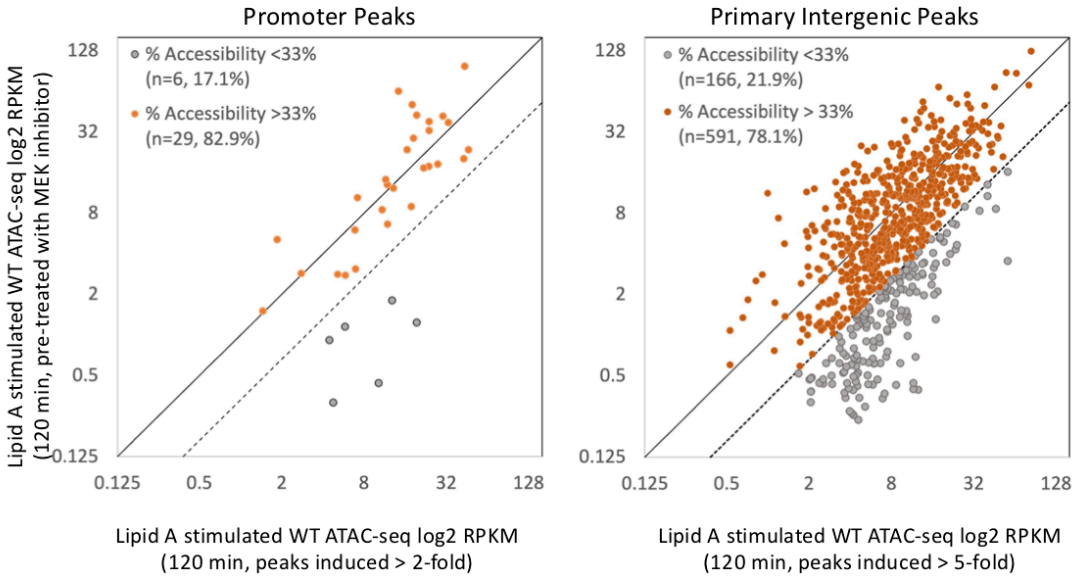
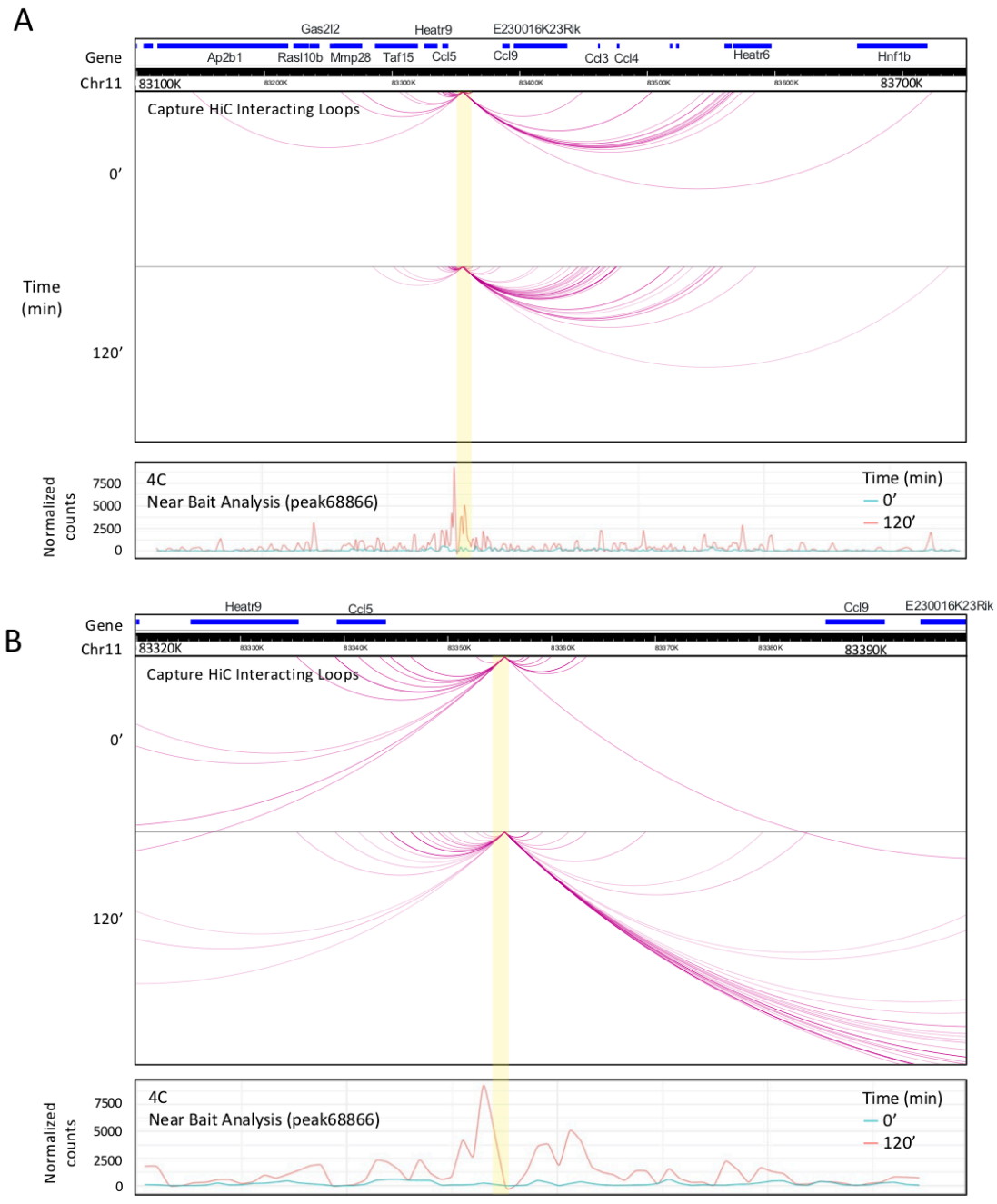


Figure B-S5. Zoomed-Out and Zoomed-In Views of 4C and Capture Hi-C Interactions with the IRF3-Dependent Intergenic ATAC-seq Peak Region Upstream of the *Ccl5* Gene



WORKS CITED

Agalioti, T., Lomvardes, S. Parekh, B., Yie, J., Maniatis, T., and Thanos, D. (2000). Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. Cell *103*, 667-678.

Barozzi, I., Simonatto, M., Bonifacio, S., Yang, L., Rohs, R., Ghisletti, S., and Natoli, G. (2014). Coregulation of transcription factor binding and nucleosome occupancy through DNA features of mammalian enhancers. Mol. Cell *54*, 844-857.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins, and nucleosome position. Nat. Methods *10*, 1213-1218. *https://doi.org/10.1038/nmeth.2688*

Cheng, Q.J., Ohta, S., Sheu, K.M, Spreafico, R., Adelaja, A., Taylor, B., and Hoffmann, A. (2021). NF-κB dynamics determine the stimulus specificity of epigenomic reprogramming in macrophages. Science *372*, 1349-1353.

Clapier, C.R., Iwasa, J., Cairns, B.R., and Peterson, C.L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nat. Rev. Mol. Cell. Biol. *18*, 407-422. *https://doi.org/10.1038/nrm.2017.26*

Comoglio, F., Simonatto, M., Polletti, S., Liu, X., Smale, S.T., Barozzi, I., and Natoli, G. (2019). Dissection of acute stimulus-inducible nucleosome remodeling in mammalian cells. Genes Dev. *33*, 1159-1174.

Frederick, M.A., Williamson, K.E., Garcia, M.F., Ferretti, M.B., McCarthy, R.L., Donahue, G., Monteiro, E.L., Takenaka, N., Reynaga, J., Kadoch, C., et al. (2023). A pioneer factor locally opens compacted chromatin to enable targeted ATP-dependent nucleosome remodeling. Nat. Struct. Mol. Biol. *30*, 31-27.

Gao, F., Elliott, N.J., Ho, J., Sharp, A., Shokhirev, M.N., and Hargreaves, D.C. (2019). Heterozygous mutations in SMARCA2 reprogram the enhancer landscape by global retargeting of SMARCA4. Mol. Cell *75*, 891-904.

Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.-L., et al. (2010). Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. Immunity *32*, 317–328.

Glass, C.K., and Natoli, G. (2016). Molecular control of activation and priming in macrophages. Nat. Immunol. *17*, 26-33. *https://doi.org/10.1038/ni.3306*

Gatchalian, J., Liao, J., Maxwell, M.B., and Hargreaves, D.C. (2020). Control of stimulus-dependent responses in macrophages by SWI/SNF chromatin remodeling complexes. Trends Immunol. *41*, 126-140.

Hargreaves, D.C. (2021). Chromatin openness requires continuous SWI/SNF activity. Nat. Genet. *53*, 263-264. *https://doi.org/10.1038/s41588-021-00781-7*

Hargreaves, D.C., Horng, T., and Medzhitov, R. (2009). Control of inducible gene expression by

signal-dependent transcriptional elongation. Cell *138*, 129–145.

Hargreaves, D.C., and Crabtree, G.R. (2011). ATP-dependent chromatin remodeling; genetics, genomics and mechanisms. Cell Res. *21*, 396-420.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Heinz, S., Romanoski, C.E., Benner, C., and Glass, C.K. (2015). The selection and function of cell type-specific enhancers. Nat. Rev. Mol. Cell. Biol. *16*, 144-154.

Hoebe, K., Du, X., Georgel, P., Janssen, E., Tabeta, K., Kim, S.O., Goode, J., Lin, P., Mann, N., Mudd, S., et al. (2003). Identification of Lps2 as a key transducer of MyD88-independent TIR signalling. Nature *424*, 743-748.

Kaikkonen, M.U., Spann, N.J., Heinz, S., Romanoski, C.E., Allison, K.A., Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C., et al. (2013). Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. Mol. Cell *51*, 310-325. *https://doi.org/10.1016/j.molcel.2013.07.010*

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, Z., Schulz, M.H., Look, T., Begemann, M., Zenke, M., and Costa, I.G. (2019). Identification of transcription factor binding sites using ATAC-seq. Genome Biol. *20*, 45.

Liu, S., Cai, X., Wu, J., Cong, Q., Chen, X., Li, T., Du, F., Ren, J., Wu, Y.-T., Grishin, N.V., et al. (2015). Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. Science *347*, aaa2630.

Liu, H., Kang, H., Liu, R., Chen, X., and Zhao, K. (2002). Maximal induction of a subset of interferon target genes requires the chromatin-remodeling activity of the BAF complex. Mol. Cell. Biol. *22*, 6471-6479.

Lone, I.N., Shukla, M.S., Richard, J.L.C., Peshev, Z.Y., Dimitrov, S., and Angelov, D. (2013). Binding of NF-κB to nucleosomes: effect of translational positioning, nucleosome remodeling, and linker histone H1. PLoS Genet. *9*, e1003830. *https://doi.org/10.1371/journal.pgen.1003830*

Martínez de Paz, A., and Josefowicz, S.Z. (2021). Signaling-to-chromatin pathways in the immune system. Immunol. Rev. *300*, 37-53.

Monticelli, S., and Natoli, G. (2017). Transcriptional determination and functional specificity of myeloid cells: making sense of diversity. Nat. Rev. Immunol. *17*, 595-607. *https://doi.org/10.1038/nri.2017.51*

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods *5*, 621–628.

Natoli, G., and Ostuni, R. (2019). Adaptation and memory in immune responses. Nat. Immunol. *20*, 783-792. *https://doi.org/10.1038/s41590-019-0399-9*

Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., Curina, A., Prosperini, E., Ghisletti, S., and Natoli, G. (2013). Latent enhancers activated by stimulation in differentiated cells. Cell *152*, 157–171.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. Cell *138*, 114–128. *https://doi.org/10/1016/j.cell.2009.04.020*

Ramirez-Carrozzi, V.R., Nazarian, A.A., Li, C.C., Gore, S.L., Sridharan, R., Imbalzano, A.N., and Smale, S.T. (2006). Selective and antagonistic functions of SWI/SNF and Mi-2beta nucleosome remodeling complexes during an inflammatory response. Genes Dev. *20*, 282-296. *https://doi.org/10.1101/gad.1383206*

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrated genomics viewer. Nat. Biotechnol. *29*, 24-26.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Sheu, K.,M., and Hoffmann, A. (2022). Functional hallmarks of healthy macrophage responses: their regulatory basis and disease relevance. Annu. Rev. Immunol. *40*, 295-321.

Siggers, T., Chang, A.B., Teixeira, A., Wong, D., Williams, K.J., Ahmed, B., Ragoussis, J., Udalova, I.A., Smale, S.T., and Bulyk, M.L. (2012). Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-κB family DNA binding. Nat. Immunol. *13*, 95–102.

Stormberg, T., Filliaux, S., Baughman, H.E.R., Komives, E.A., and Lyubchenko, Y.L. (2021). Transcription factor NF-κB unravels nuclesomes. Biochim. Biophys. Acta Gen. Subj. *1865*, 129934.

Tong, A.J., Liu, X., Thomas, B.J., Lissner, M.M., Baker, M.R., Senagolage, M.D., Allred, A.L., Barish, G.D., and Smale, S.T. (2016). A stringent systems approach uncovers gene-specific mechanisms regulating inflammation. Cell *165*, 165-179. *https://doi.org/10.1016/j.cell.2016.01.020*

Tamura, T., Yanai, H., Savitsky, D., and Taniguchi, T. (2008). The IRF family transcription factors in immunity and oncogenesis. Annu. Rev. Immunol. *26*, 535–584.

Toshchakov, V., Jones, B.W., Perera, P.-Y., Thomas, K., Cody, M.J., Zhang, S., Williams,

B.R.G., Major, J., Hamilton, T.A., Fenton, M.J., et al. (2002). TLR4, but not TLR2, mediates IFN-beta-induced STAT1alpha/beta-dependent gene expression in macrophages. Nat. Immunol. *3*, 392–398.

Vierbuchen, T., Ling, E., Cowley, C.J., Couch, C.H., Wang, X., Harmin, D.A., Roberts, C.W.M., and Greenberg, M.E. (2017). AP-1 transcription factors and the BAF complex mediate signal-dependent enhancer selection. Mol. Cell *68*, 1067-1082.

Weinmann, A.S., Plevy, S.E., and Smale, S.T. (1999). Rapid and selective remodeling of a positioned nucleosome during the induction of IL-12 p40 transcription. Immunity *11*, 665-675.

Wolf, B.K., Zhao, Y., McCray, A., Hawk, W.H., Deary, L.T., Sugiarto, N.W., LaCroix, I.S., Gerber, S.A., Cheng, C., and Wang, X. (2023). Cooperation of chromatin remodeling SWI/SNF complex and pioneer factor AP-1 shapes 3D enhancer landscapes. Nat. Struct. Mol. Biol. 30, 10-21.

Yamamoto, M., Sato, S., Hemmi, H., Hoshino, K., Kaisho, T., Sanjo, H., Takeuchi, O. Sugiyama, M., Okabe, M, Takeda, K., et al. (2003). Role of adaptor TRIF in the MyD88-independent Toll-like receptor signaling pathway. Science *301*, 640-643.

Zambelli, F., Pesole, G., and Pavesi, G. (2009). Pscan: finding over-represented transcription factor binding site motifs in sequences from co-regulated or co-expressed genes. Nucleic Acids Res. *37*, W247–W252.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. 9, R137.