UNIVERSITY OF CALIFORNIA, SAN DIEGO

Innate immune gene expression programs defined by both regulated mRNA synthesis and decay

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

in

Bioinformatics and Systems Biology

by

Christine Shiang-Ling Cheng

Committee in charge:

Professor Alexander Hoffmann, Chair
Professor Charles Elkan, Co-Chair
Professor Christopher K. Glass
Professor Bing Ren
Professor Shankar Subramaniam

2011

The Dissertation of Christine Shiang-Ling Cheng is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____
co-Chair

_____
Chair

University of California, San Diego

2011

DEDICATION

I dedicate this thesis to my father, 鄭國順, who as a Professor of Mathematics has greatly influenced me to become interested in research and science. He always has strong confidence on my ability and is extremely supportive of me no matter what I have decided to pursuit. I also would like to dedicate this thesis to my mother, 葉錦鳳, who together with my father have raised my five years old son for me so that I can focus on getting my degree. I owe them so much that there is no way I can return in equal. I also dedicate this thesis to my husband, 張昶 (Chang Chang), who has been my biggest support since we know each other 15 years ago.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Alex Hoffmann, for providing me with tremendous supports when I am exploring unknown or risky territories but at the same time constantly pull me back from unrealistic or overly ambitious plans and attempts. He has taught me how to structure and frame a study with a more realistic scope so that it becomes possible to actually complete the study. His enthusiasm for science had great impact on me and has been the driving force for many of the interesting ideas that came out of our scientific discussion. I would also like to thank my previous mentor, Geoff Rosenfeld, for providing me with the training to become capable of performing biological experiments. He has also greatly influenced me by many forward thinking concepts that have help shaped my understanding of the biological system. I would also like to thank my committee, Bing Ren, Chris Glass, Charles Elkan and Shankar Subramaniam for their insightful suggestions and great support throughout my graduate career.

I would like to thank James Lee for sharing some crazy graduate school days in the lab and outside the lab and for being a great friend. I would also like to thank Masa Asagiri for his insightful suggestions on my study and manuscript, his suggestions for my career development and for his companies at late nights. I also thank Paul Loriaux for suggestions on my manuscript and for his witty humor that always make me laugh and for his friendship. I thank Chris Benner for letting me use his motif search program

Homer and his suggestions and support. I thank Anna Krones from the Rosefeld lab for helping me debug through my ChIP assays and for her support and being a great friend. I also thank Aakash Patel for providing excellent help with experiments and for giving me a chance to learn how to mentor a student.

I would like to thank my parents for their support and believe in me. I would also like to thank my husband for his understanding, support, and always being there for me. I also would like to thank him for the scientific discussions and his technical help with mathematical models.

Chapter 3, in full, is a reprint of the material as it appears in Science Signaling 2011**.** Cheng CS, Feldman KE, Lee J, Verma S, Huang DB, Huynh K, Chang M, Ponomarenko JC, Sun SC, Benedict CA, Ghosh G, Hoffmann A, American Association for the Advancement of Science (AAAS), 2011. The thesis author was the primary investigator and author of this paper.

VITA

1996 Bachelor of Science, Botany, National Taiwan University

2001 Master of Science, Computer Science, Stanford University

2011 Doctor of Philosophy, Bioinformatics and Systems Biology, University of California, San Diego

PUBLICATIONS

Escoubet-Lozach L, Benner C, Kaikkonen MU, Lozach J, Heinz S, Spann NJ, Crotti A, Stender J, Ghisletti S, Reichart D, **Cheng CS**, Luna R, Eckhardt C, Sasik R, Garcia-Bassets I, Hoffmann A, Subramaniam S, Hardiman G, Rosenfeld MG, Glass CK. Mechanisms Establishing TLR4-Responsive Activation States of Inflammatory Response Genes. *Submitted to PLOS Genetics, in revision*.

Cheng CS, Feldman KE, Lee J, Verma S, Huang DB, Huynh K, Chang M, Ponomarenko JC, Sun SC, Benedict CA, Ghosh G, Hoffmann A. The specificity of innate immune responses is enforced by repression of interferon response elements by NFκB p50. Science Signaling. 2011 Feb 22;4(161):ra11.

Ramirez-Carrozzi VR, Braas D, Bhatt DM, **Cheng CS**, Hong C, Doty KR, Black JC, Hoffmann A, Carey M, Smale ST. A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. Cell. 2009 Jul 10;138(1):114-28.

Cheng CS, Johnson TL, Hoffmann A. Epigenetic control: slow and global, nimble and local. Genes and Development. 2008 May 1;22(9):1110-4. Perspective.

Garcia-Bassets I, Kwon YS, Telese F, Prefontaine GG, Hutt KR, **Cheng CS**, Ju BG, Ohgi KA, Wang J, Escoubet-Lozach L, Rose DW, Glass CK, Fu XD, Rosenfeld MG. Histone methylation-dependent mechanisms impose ligand dependency for gene activation by nuclear receptors. Cell. 2007 Feb 9;128(3):505-18.

Kwon YS, Garcia-Bassets I, Hutt KR, **Cheng CS**, Jin M, Liu D, Benner C, Wang D, Ye Z, Bibikova M, Fan JB, Duan L, Glass CK, Rosenfeld MG, Fu XD. Sensitive ChIP-DSL technology reveals an extensive estrogen receptor alpha-binding program on human gene promoters. Proc Natl Acad Sci U S A. 2007 Mar 20;104(12):4852-7.

ABSTRACT OF THE DISSERTATION

Innate immune gene expression programs defined by both regulated mRNA synthesis and decay

by

Christine Shiang-Ling Cheng

Doctor of Philosophy

University of California, San Diego, 2011

Professor Alexander Hoffmann, Chair

Professor Charles Elkan, Co-Chair

The innate immune system elicits a complex pathogen-specific inflammatory gene expression program involving hundreds of genes to provide the first line of host defense against pathogens. Two major transcription factor families, nuclear factor κB (NFκB) and the interferon regulatory factors (IRFs) are known to bind the κB-site and the interferon regulatory element (IRE), respectively. To our surprise, we identified the NFκB p50 homodimer as a negative regulator of IRF and anti-viral responses by directly binding to the IRE containing promoters via a newly defined subclass of guanine-rich IRE (G-IRE). Furthermore, we found the expression of the antiviral regulator IFNβ to be

stimulus-restricted by p50 homodimer binding to the G-IRE-containing enhancer to suppress cytotoxic IFN signaling.

The inflammatory expression program is believed to be a complicated gene regulatory network that involves more than 100 transcriptional regulators. Utilizing defined cellular stimuli and mathematical modeling of stereotypical promoter architectures, we categorized the inflammatory gene activation response into three surprisingly simple and separable gene programs that are functional targets of NFκB, bZIP and IRF/ISGF3. We did not identify a class of genes whose expression depends on synergy between different transcription factors, however, we discovered a new class of pathogen specific gene expression which depends on synergy between regulated mRNA stability and transcription.

MicroRNA-mediated gene repression has emerged as a potent mechanism of biological regulation pertaining to cancer and immunity, yet identifying target genes of specific miRNAs has been difficult. We pursued a computational discovery strategy that is inclusive of post-transcriptional modifications to miRNA seed sequences. We identified a specific A-to-I edited miR-155 seed sequence to be enriched in the 3'UTR of gene repression programs in macrophages and inflammatory TH1 cells, but not in anti-inflammatory TH2 cells or non-immune cells. In fact, edited forms of many highly expressed miRNAs emerge as candidate gene repression factors, in correlation with the strength of IFN signaling and the expression of an interferon inducible adenosine deaminase, ADAR1.

Chapter 1 : Introduction

The innate immune system is the first line of host defense against pathogens. The initial sensing of pathogen-associated molecular patterns (PAMPs) are by the germline-encoded pattern recognition receptors (PRRs). More recently, evidence has emerged that the PRRs also recognize endogenous molecules, damage-associated molecular patterns (DAMPs) that are released from damaged cells (Takeuchi and Akira, 2010). Toll-like receptors (TLRs) are the most well studied PRRs. They are differentially expressed on various immune cells, including dendritic cells, macrophages, B cells and specific types of T cells. Each TLR recognizes a distinct set of microbial components, TLR4 recognizes the gram-negative cell wall component lipopolysaccharide (LPS), TLR9 recognizes to the bacterial CpG-rich DNA sensor, TLR3 recognizes double stranded RNA viruses while TLR7 and TLR9 recognizes single stranded RNA viruses (Akira et al., 2006).

The signaling pathways activated by the PRRs induce the nuclear translocation of a set of transcription factors, including NFκB, IRFs and AP-1, which leads to the activation of hundreds of inflammatory response genes, such as cytokines, chemokines, IFNs, genes that are involved in antimicrobial defense and tissue repair (Takeuchi and Akira, 2010). It is important to understand the regulatory network that determine the innate immune response and inflammatory gene expression program, as uncontrolled activation of the inflammatory response could contribute to many chronic diseases such arthritis, atherosclerosis, autoimmune diseases or cancer (Medzhitov, 2008; Medzhitov and Horng, 2009). Therapeutic manipulation of the innate immune response could

potentially be developed into effective innate immune vaccines or immunologic adjuvants that provide protection from invading microbes.

The pathogen elucidated gene expression program in macrophage and dendritic cells can be broadly classified into antiviral functions and inflammatory responses (Amit et al., 2009; Foster et al., 2007). Many studies indicate that individual genes are regulated in a gene specific manner through the combinatorial control of pathogen activated transcription factors (Litvak et al., 2009) or chromatin modifications (Foster et al., 2007). However, systematic knockdown of 125 transcriptional regulators in dendritic cells uncovered that there are still only two classes of transcription factors among the 125 regulators, the antiviral response regulators, and the inflammatory response regulators (Amit et al., 2009). The induction profiles of mRNA expression levels may not only be a function of synthesis (transcription) but are also of degradation (mRNA halflife). The temporal gene activation profiles induced by the inflammatory cytokine, TNF, correlates with the mRNA halflife, such that early response genes have short mRNA halflives and late response genes have longer mRNA halflives (Hao and Baltimore, 2009). It is also interesting to note that the induction of some of these early response genes does not require chromatin remodeling events, while the induction of the late response genes requires chromatin remodeling event to occur before the gene can be activated(Ramirez-Carrozzi et al., 2009). An integrated view of chromatin modification, chromatin remodeling, transcriptional activation and mRNA stability events are needed to understand the inflammatory gene expression program.

The recent advancement in high-throughput technologies, such as chromatin immunoprecipitation followed by sequencing (ChIP-seq), RNA profiling by sequencing (RNA-seq) and genome-wide association (GWA) studies has drastically enhanced our ability to characterize the biological system in a top down approach(Hawkins et al.). In the genome-wide era, analysis approaches become the critical tool that allows us to "observe" and come up with biological insights and hypothesis. Bottom up approaches, such as mathematical modeling provide a cost effective way to test multiple hypothesis and potentially pin point the most promising hypothesis through simulations (Kim et al., 2009). In my thesis work, I applied top down and bottom up approaches with both experimental and computational techniques in an integrated way to dissect the inflammatory gene expression program.

I aimed to characterize the inflammatory expression program by performing microarray studies of primary macrophages and primary mouse embryonic fibroblast cells from mice deficient in different transcriptional factors or different signaling pathways. We also utilized a bottom-up approach by building mathematical models that allow us to rapidly iterate through different regulatory mechanisms that can potentially explain the gene expression profiles that we obtained from the top-down approach. From combing the top-down and bottom-up approaches we have identified novel regulatory networks for the NFκB p50 homodimer, which to our surprise, functions to regulate the interferon responsive genes that contain guanine-rich interferon response elements (G-IREs) rather than NFκB target genes. We also categorized the inflammatory expression response into three simple and separable gene programs that are functional targets of

NFκB, bZIP and IRF/ISGF. Furthermore, we have identified a new class of gene expression that controlled by the synergy between regulated mRNA decay and transcription. More interestingly, by employing a computational strategy that considers edited miRNA seed sequences, we found a specific A-to-I edited form of the miR-155 seed sequence is highly enriched in the 3'UTRs of gene repression programs in activated immune cells, the enrichment are correlated with the strength of the IFN pathway and the expression of the deaminase, ADAR1.

Chapter 2 : Synergy between regulated mRNA synthesis and decay defines the cellular response to pathogen exposure

ABSTRACT

The innate immune system elicits a complex pathogen-specific inflammatory gene expression program involving hundreds of genes to provide a first line of host defense against pathogens. It is important to understand the gene regulatory networks that determine the response to pathogen-associated molecular patterns (PAMPs). We probed the system with agents that stimulate a subset of the signaling systems, employed mathematical models of in silico gene circuits, and categorized the gene activation response into three separable gene programs that are functional targets of NFκB, bZIP and IRF/ISGF. Bioinformatic regulatory motif analyses and mouse genetics further validated our model prediction of three separable programs. We did not identify a class of genes whose expression depends on synergy between different transcription factors, however, we discovered a new class of PAMP-specific gene expression which depends on synergy between regulated mRNA stability and transcription. Our results suggest that the synergistic combinations of two distinct regulatory mechanisms – synthesis and degradation – produce the unusually potent cellular gene expression response to pathogens.

INTRODUCTION

Pathogen recognition by the innate immune system is a specific and tightly orchestrated process that involves the recognition of pathogen-associated molecular patterns (PAMPs) by the germline-encoded pattern recognition receptors (PRRs) such as the Toll-like receptors (TLRs), the Retinoic acid-inducible gene (RIG)-I-like receptors (RLRs) and NOD-like receptor (NLRs)(Kawai and Akira; Takeuchi and Akira). These PRRs are expressed not only in immune cells, such as macrophages and dendritic cells, but are also expressed in nonprofessional immune cells including fibroblasts and epithelial cells(Takeuchi and Akira). The intra-cellular signaling pathways triggered by the PRRs induce an inflammatory gene expression program that involves the activation of hundreds of genes, including cytokines, chemokines, IFNs and genes involved in antimicrobial defense and tissue repair(Takeuchi and Akira). The inflammatory response is beneficial as it provides protection against microbial infection or tissue damage, but it can also be detrimental when missregulated(Medzhitov, 2008). Uncontrolled activation of the inflammatory response can be fatal or may cause or contribute to many chronic diseases such arthritis, atherosclerosis, autoimmune diseases or cancer.

The family of TLR receptors, ten in humans and 12 in mice, each recognize various components from bacteria, parasites and viruses and then differentially activates specific cellular signaling pathways that are required to fight the specific pathogen(Kawai and Akira; Takeuchi and Akira). Bacterial lipopolysaccharide (LPS), one of the most potent TLR agonist, is recognized by the TLR4 receptor, which engages both of the two

master signaling adaptors, MyD88 and TRIF; Myd88 is known to activate the mitogen-activated protein kinases (MAPKs) and early NFκB, whereas TRIF subsequently activates IRF3 and late NFκB(Kawai and Akira; Takeuchi and Akira) (Fig. 1A). These events activate two autocrine regulators: the important anti-viral regulator IFNβ activates the Stat1-, Stat2- and IRF9-containing ISGF3 complex inducing the expression of a large class of anti-viral genes that contain interferon regulated elements (IREs) in their promoter regions(Borden et al., 2007; Doyle et al., 2002); LPS also induces the expression of PDGFβ which amplifies MAPK signaling to the bZIP-transcription factors AP-1, ATF, CREB(Chow et al., 2005).

Given its physiological relevance, several previous studies have attempted to understand the underlying transcriptional regulatory network. Comprehensive microarray analyses revealed the gene expression response to different TLR ligands to be not only remarkably extensive involving more than 1000 genes, but also potentially very complex as the expression of more than 100 transcription factors was shown to be regulated during the inflammatory response(Ramsey et al., 2008). Indeed, specific target genes were shown to be regulated by induced expression of the negative regulator ATF3[7,] or the positive regulator C/EBPδ(Gilchrist et al., 2006; Litvak et al., 2009). However, systematic single knockdown of 125 transcriptional regulators in dendritic cells found that few of these candidate regulators were required(Amit et al., 2009), suggesting that either single regulators are redundant with others or that the transcriptional network is in fact not as complex as previously suggested. To distinguish between these possibilities and dissect the gene regulatory circuitry of the LPS-induced expression response, we

pursued a synthetic approach using both experimental and mathematical modeling strategies. To this end we employed cellular cytokine and growth factor stimuli that engage a defined set of signaling pathways and constructed mathematical models of combinatorial promoter architectures to determine how much of the LPS-gene expression response could be accounted for. We found that three primary transcription factor systems (bZIP, NFκB, ISGF3) functioning single or additively, but not synergistically could account for 85% of the response. A major LPS-specific expression program turned out not to be the result of synergy of transcription factors but, unexpectedly, of synergy between mechanisms regulating mRNA synthesis via NFκB activation and mRNA stability via control of the RNA binding protein tristetraproline (TTP).

RESULTS

Dissect the pathogen response by using defined cellular stimuli

In the cellular response to LPS, more than 100 transcription factors have been implicated and are thought to function coordinately, providing feedback, and/or feedforward loops(Amit et al., 2009; Ramsey et al., 2008). To dissect this complex gene expression response, we selected cellular stimuli that are known to activate a defined subset of the LPS induced signaling pathways. Platelet-derived growth factor B (PDGFβ), the major growth factor in blood serum, is known to activate the MAPK pathway(Tallquist and Kazlauskas, 2004). Tumor necrosis factor (TNF), a pleiotropic inflammatory cytokine that regulates diverse cellular responses which include cell death,

survival, proliferation, differentiation and migration, is known to activate the MAPK and NFκB pathways(Bradley, 2008). IFNβ, the important antiviral autocrine protein that has been widely used for treating viral disease, multiple sclerosis and cancer, is known to activate the ISGF3 transcription factor(Borden et al., 2007). A schematic showing the downstream signaling and transcription factor pathways activated by each stimulus is shown in Figure 1A.

Consistent with the literature, while LPS activates all pathways, we observed activation of the NFκB pathway only in response to TNF, activation of the ISGF3 pathway only in response to IFNβ, and MAPK activation by both TNF and PDGFβ but not IFNβ (Figure 1B and Figure 1C). To characterize the expression programs responsive to each stimulus, we performed microarray analysis of gene expression in response to LPS, TNF, PDGFβ and IFNβ in primary mouse embryonic fibroblast cells (MEFs). We performed analysis with Gene Expression Dynamic Inspector that uses a self-organizing map (SOM) to reduce data dimensionality and to create characteristic visual representation of each sample(Eichler et al., 2003). We identified PDGFβ specific genes that are induced by PDGFβ only at 1 hour but is induced by TNF at 1, 3 and 8 hours (Figure 1D, orange circles). TNF induces a specific group of genes at 3 and 8 hour (Figure 1D, red circles) and IFNβ induces a distinct set of genes (Figure 1D, green circles). Strikingly, gene expression profiles from the 3 cellular stimuli added up to constitute the LPS induced expression program, with even the temporal dynamics from

each stimulus being conserved. However, the SOM analysis also identified a group of genes specifically induced by LPS,but are not induced by the three cellular stimuli.

To further characterize stimulus specific expression profiles, we performed K-means clustering analysis that provides further levels of detail. Similar to what we observed with SOMs, K-means clustering also identified three distinct classes of genes (though split into several clusters) that are PDGFβ specific (Figure 2A, cluster A and B), TNF specific (Figure 2A, cluster C and D) and IFNβ specific (Figure 2A, cluster F, G and H) and a cluster E that is only induced by LPS. Interestingly, the two autocrine mechanisms activate their effectors as expected: IFNβ induction is followed byISGF3 activation and PDGFβ induction is followed by a second phase of MPK activity (Figure 1C). We also performed microarray analysis with the same set of stimuli in bone marrow derived macrophages (BMDMs) and fetal liver derived macrophages (FLDMs) and have also observed the three distinct stimulus specific classes of genes (Data not shown).

Identify prevalent promoter architectures with mathematical models

To infer potential transcriptional regulatory circuits underlying the stimulus-specific gene expression programs, we constructed mathemcatical models of a small set of stereotypical promoter architectures (Figure 3A). Thermodynamic expressions calculate the probability of promoter occupancy as a function of transcription factor binding activity on the specific binding site(Bintu et al., 2005a; Bintu et al., 2005b). To calculate mRNA expression level of various in silico genes, the thermodynamic

formulation of promoter activities is imbeded in a differential equation describing mRNA production and decay.

Using promoters with both an IRE and a κB binding site as an example, an AND gate promoter requires the activation of both ISGF3 and NFκB to induce transcriptional activation (Figure 3B). The IFNb enhancer is a classical example of such an AND gate promoter where coordinated activation of AP1, NFκB and IRF are required to synergistically activate gene expression(Escalante et al., 2007; Thanos and Maniatis, 1995). An OR gate promoter is defined as activation of either one of the transcription factors is sufficient to induce transcriptional activation though the two may also function additively (Figure 3C).

We considered promoter architectures that involve three transcriptional activators, NFκB, bZIP and ISGF3, either singly, or doubly or triply to form combinatorial AND or OR gates, which resulted in 11 distinct promoter architectures (Figure 4A). We also considered short and long mRNA half-lives as mRNA half-life control is known to affect the temporal pattern of expression. Activity profiles of transcription factors or upstream signaling kinases in response to each stimulus (Figure 1B and Figure 1C) were used as inputs to the models to generate stimulus specific in silico expression profiles. We present our in silico expression profiles as gene expression heatmaps produced by the 33 in silico genes (Figure 4A).

Matching the in silico gene expressio profiles with the in vivo gene expression profiles, our model predicted that 75% (Clusters A, B, C, D, G and H from Figure 2) of the stimulus specific gene expression can be explained by single transcription factor regulation, such as in silico NFκB target (gene 1), bZIP target (gene 2) and ISGF3 target (gene 3) (Figure 4B). As expected, our models suggested that mRNA half-lives plays a role in determining the temporal expression profiles. For example, the NFκB target genes in cluster C are predicted to have shorter mRNA half-lives as compared to NFκB targets in cluster D, indeed, mRNA half-lives measured by actinomycin D treatment show shorter half-lives for genes in cluster C as compared to cluster D (Figure 4B and 4C). Similarly, ISGF3 target genes in cluster G have shorter mRNA half-lives as compared to genes in cluster H (Figure 4B and 4C). Thus, our mathematical models indicate that a large portion of LPS-responsive gene expression can be explained by three classes of single transcription factor regulated promoters without considering transcriptional AND or OR gates. In addition, the mathematical models identified genes in cluster F as potential OR gates between NFκB and ISGF3 (Fig. 4B). Cluster E, on the other hand is LPS-specific and matched by in silico AND-gate promoters on which NFκB and ISGF function synergistically.

Bioinformatic and experimental analyses of model predictions

To validate our model prediction, we performed motif analyses with promoter sequences that are 1000 bp upstream and 300 bp downstream of the transcriptional start sites. We observed enrichment of the bZIP binding sites in the predicted bZIP target

gene, cluster B, though not in cluster A (Figure 5A). We also observed enrichment of the NFκB binding sites in the promoters of clusters C and D, the two NFκB target clusters predicted from our mathematical models (Figure 5A). We also observed enrichment of the IRE binding sites in the promoters of the predicted ISGF3 target genes, clusters G and H. Interestinly, cluster F (OR gate cluster) is enriched for both the NFκB and the IRE binding sites, as predicted. However, contrary to prediction cluster E (the predicted AND gate cluster) is enriched for only the NFκB binding site (Figure 5A).

Our modeling prediction of three distinct classes of genes was further confirmed with microarray analysis of MEFs from mice deficient in NFκB ($crel^{-/-}p65^{-/-}$), ISGF3 ($ifnar^{-/-}$) and both NFκB and ISGF3 ($crel^{-/-}p65^{-/-}irf3^{-/-}$). LPS induced activation of NFκB target genes (clusters C, D and E) were abolished in $crel^{-/-}p65^{-/-}$ MEFs, and the OR gate genes (cluster F) were partially effected in the $crel^{-/-}p65^{-/-}$ cells, while ISGF3 targets (clusters G and H) were unaffected (Figure 5B). On the contrary, LPS induced activation of the ISGF3 target genes (clusters G and H) were dimished in the $ifnar^{-/-}$ MEFs, and the OR gate genes (cluster F) were partially effected in the $ifnar^{-/-}$ cells, while NFκB target genes (clusters C, D and E) were unaffected (Figure 5B). More interestingly, LPS did not induce any gene activation in the $crel^{-/-}p65^{-/-}irf3^{-/-}$ MEFs where both NFκB and ISGF3 were deficient. Abolished LPS-induced PDGFβ expression in the $crel^{-/-}p65^{-/-}irf3^{-/-}$ MEFs may account for the lack of expression of clusters A and B. Our results validated our model prediction that PDGFβ specific genes are bZIP target genes, TNF specific genes are NFκB target genes , IFNβ specific genes are ISGF3 target genes. Interestingly, the

OR gate genes (cluster F) allow for independent contributions by both NFκB and ISGF3. Thus, our combined experimental/modeling analysis indicates that LPS-responsive gene expression can largely be accounted for by the sum of three transcription factor axes: bZIP, NFκB and ISGF. However, the AND gate genes (cluster E) only showed over-represented κB but not IRE sites, suggesting the possibility that a fourth signaling axis that is LPS specific – not considered in our simple modeling framework – may critically contribute to their regulation.

To examined these preliminary conclusions experimentally, we exposed cells to both TNF and IFNβ to activate all three transcription factor axes with defined cellular stimuli. Interestingly, co-stimulation of TNF and IFNβ did not produce stronger gene induction than LPS, TNF or IFNb alone for all clusters from microarray analyses (Figure 6A). More specificly, cluster F represents an OR gate between NFκB and ISGF3 where the promoters are enriched for both κB and IRE binding sites (Figure 5A), while gene induction in response to TNF or IFNβ add up to gene induction with LPS, and TNF, IFNβ co-stimulation induced the same level of expression as LPS alone (Figure 6A, cluster F). Indeed, LPS induced gene activation from NFκB ($crel^{-/-}p65^{-/-}$) or ISGF3 ($ifnar^{-/-}$) deficient MEFs was partially abolished (Figure 6C). Furthermore, the only potential AND gate cluster (cluster E) are only enriched for NFκB binding sites at the promoters (Figure 5A) and the LPS induced expression are dependent on NFκB (abolished in $crel^{-/-}p65^{-/-}$), but not so much on ISGF3 ($ifnar^{-/-}$) (Figure 6B), which suggest NFκB work together like an AND gate with another transcription factor, bZIP, ISGF3 or factor X on

the cluster E promoters (Figure 6A, cluster E). Since cluster E is not induced with TNF alone, we ruled out the possibility of an AND gate between NFκB and bZIP (Figure 6A, cluster E). Co-stimulation of TNF and IFNβ did not induce much more gene activation, indicating there is no synergy between NFκB and ISGF3 (Figure 6A, cluster E). These data suggest that the only possible transcriptional regulatory circuit for cluster E is a transcriptional AND gate between NFκB and a transcription factor X that is induced only by LPS stimulation.

AND gate genes are regulated by mRNA stability

To explore the possibility that a transcription factor X forms an AND gate with NFκB at the promoter of cluster E genes and induce gene activation in response to LPS but not TNF, we measured nascent transcript levels as described previously(Giorgetti et al.). To our surprise, we found that while the mature mRNAs were strongly induced by LPS but not TNF (Figure 7A), nascent transcript levels were equally induced by LPS and TNF (Figure 7B). Our results rule out the possibility that these genes are regulated by a transcriptional AND gate in response to LPS, but instead suggest a post-transcriptional regulatory mechanism. We tested the possibility that the regulation is at the mRNA stability level and measured the mRNA half-lives by inhibiting transcription with actinomycin D. Interestingly, these mRNAs have very short half-lives in basal state or in response to TNF but LPS stimulation stabilized these mRNAs (Figure 7C).

LPS has been shown to induce mRNA stabilization through the activation of p-38 and ERK and the subsequent phosphorylation of the ARE binding protein TTP(Brook et al., 2006; Mahtani et al., 2001). Phosphorylation of TTP inhibited its interaction with the basic RNA decay machinary and thus inhibited TTP's ability to rapidly degrade mRNAs(Sandler and Stoecklin, 2008). Indeed, LPS does induce much stronger activation profiles of p-38 and ERK as compared to TNF stimulation in BMDMs (Figure 8A). Similarly, we also observed stronger activation of p-38 and ERK stimulated with LPS as compared to TNF in MEFs (Figure 1C). LPS specific induction of the mature mRNA was attenuated in the erk2$^{-/-}$ BMDMs or in BMDMs pre-treated with p38 inhibitors (Figure 8B), while LPS induced nascent transcript activation is not affected in erk2$^{-/-}$ BMDMs or by p38 inhibitor treatment (Figure 8C). LPS induced mRNA stabilization as measured by actinomycin D treatment was attenuated in the erk2$^{-/-}$ BMDMs or by p38 inhibitors treatment (Figure 8D). Many genes in cluster E are also well known TTP target genes from the literature (Al-Souhibani et al., 2010; Emmons et al., 2008; Lai et al., 2006; Stoecklin et al., 2008). Our results provided an unconventional case of an AND gate between transcription and mRNA half-life control that determines stimulus specific expression profiles (Figure 9A).

DISCUSSION

Through a combination of top down (microarray analyses, motif search and mice genetics) and bottom up (mathematical modeling of gene circuits) approaches, we classified the inflammatory expression response into three separable gene programs that

are distinct functional targets of NFκB, bZIP and IRF/ISGF3. The surprisingly simple classification of genes can explain 75% of the stimulus specificity in the inflammatory expression program, which is contradictory to the complicated regulatory network that people have assumed or hypothesized. Interestingly, microarray study of the innate immune expression program after systematic knockdown of 125 transcriptional regulators also identified two major class of transcriptional regulators, the pro-inflammatory regulators (NFκB and IκBs) and the anti-viral regulators (IRFs and Stats)(Amit et al., 2009). Furthermore, as much as 70% of the 125 transcriptional regulators they have examined did not have much functional effect after shRNA knockdown (Fig. 3C of Amit et al.(Amit et al., 2009)), which supports our idea that the innate immune expression program is not regulated by a huge number of transcription factors, but instead is regulated by the three major class of transcription factors, NFκB, bZIP and IRF/ISGF3.

It has recently been shown that mRNA stability affects the temporal order of the induction of some inflammatory response genes, where mRNAs that are expressed early have shorter mRNA half-lives while mRNAs that are induced later have longer mRNA halfe-lives(Hao and Baltimore, 2009). Interestingly, our results extends this concept to a larger scale, we show that different temporal expression profiles within each group can be completely explained by the differences in mRNA half-lives (Fig. 5C, cluster C vs cluster D for NFκB targets and cluster G vs cluster H for ISGF3 targets). Furthermore, consistent classification of the three groups of genes in MEFs, BMDMs and FLDMs suggest our finding is a general characteristic of the inflammatory expression program. In contrast to

the transcription factor binding targets identified by chromatin immunoprecipitation or ChIP-seq experiments, we have identified functional targets that are biologically relevant. However, we do not rule out the possibility that smaller transcriptional circuits or modules might exist where the expression level of the transcription factor is induced by NFκB, bZIP or ISGF3.

Instead of using pathological agonists that target different TLR pathways, the stimuli that we have selected are endogenous ligands (TNF, PDGFβ and IFNβ) that are important for the normal physiology. It might be important for an organism to have predefined functional modules that are mutually exclusive and non-overlapping to provide needed functional distinction and the flexibility to mix and match between different modules. For example, in physiological conditions where PDGFβ as an important growth factor is playing a role, anti-viral function may not be needed and could potentially be detrimental if unintentionally induced. The NFκB pathway has many cross regulation between the anti-viral IRF/IFN pathway and many pathogen components induce the activation of both pathways, however, our findings show that the NFκB pathway is a distinct and mutually exclusive functional module from the IRF/IFN pathway (except for the OR gate genes in cluster F in Figure 5).

We did not observe any substantial number of genes that could be regulated by an AND gate between transcription factors. Instead, we discovered an unusual AND gate that is between mRNA half-life control and transcriptional regulation (Cluster E in Figure 5). The additional regulation from stimulus induced mRNA stability adds another

dimension to the combinatorial transcriptional regulatory circuits. We also speculate that stimulus induced mRNA half-life control may provide faster responses as stalled degradation of existing mature mRNAs can readily contribute to the increase of steady state mRNAs without new synthesis and this could be analyzed by mathematical modeling. Number of AREs in the 3'UTR region also correlates with the stability of the mRNA possibly through the regulation of TTP. Many of the genes in our mRNA half-life AND gate cluster are known target of TTP from the literature (Al-Souhibani et al., 2010; Emmons et al., 2008; Lai et al., 2006; Stoecklin et al., 2008). Interestingly, recent work suggests that the level of TTP is lower in cancer cells as compared to normal cell types and that many of the inflammatory response genes (also in our cluster E in Figure 5) are TTP targets in breast cancer cells(Al-Souhibani et al.). This implies that our mRNA half-live AND gate genes that are also NFκB target genes might be involved in the cancer related process and could potentially be the trigger for hyper-inflammation in cancer cells. It would be interesting to understand how the two pathways, NFκB and TTP/mRNA half-lives, cross regulate within the cancer biology.

MATERIAL AND METHODS

Cell culture.

BMDMs were isolated from C57BL/6 and *erk21*$^{-/-}$ mice and cultured in L929 cell-conditioned medium for 5-8 days. Primary MEFs were prepared with E12-E14 embryos

from C57BL/6, *ifnar*$^{-/-}$, *crel*$^{-/-}$*p65*$^{-/-}$, *crel*$^{-/-}$*p65*$^{-/-}$*irf3*$^{-/-}$ mice and cultured in DMEM containing 10% BCS for 5-6 passsages. Cells were stimulated with 0.1ug/ml LPS (Sigma, B5:055), 2500 U/ml murine IFNβ (Biogen, Inc), 10ng/ml TNF and 50 ng/ml recombinant human PDGF-BB (Roche).

Transcriptome and Bioinformatic analysis.

RNA were extracted with Qiagen RNeasy kit and hybridized to Illumina mouse RefSeq Sentrix-8 V1.1 and V2 BeadChips at the UCSD Biogem facility. De novo motif search were performed with the promoter sequences 1kb upstream and 0.3kb downstream of the transcription start site using the motif search program Homer, developed by Dr. Chris Benner (Heinz et al.). An in-depth description and benchmarking of this software suite can be found at http://biowhat.ucsd.edu/homer/.

Biochemical Assays.

Immunoblots and EMSAs were conducted using standard methods as described (Basak et al., 2007; Werner et al., 2005). Total RNA was isolated using Qiagen RNAeasy kit from BMDMs or MEFs treated as indicated. RNA was reverse transcribed with Superscript II reverse transcriptase (Invitrogen) and resulted cDNA was used for real-time qPCR analysis (SYBRgreen).

ACKNOWLEDGEMENTS

I thank Aakash Patel for performing the immunoblots for phospho-MAPKs in MEFs and BMDMs. I also thank Dr. Chris Benner for letting me use his motif search program Homer.

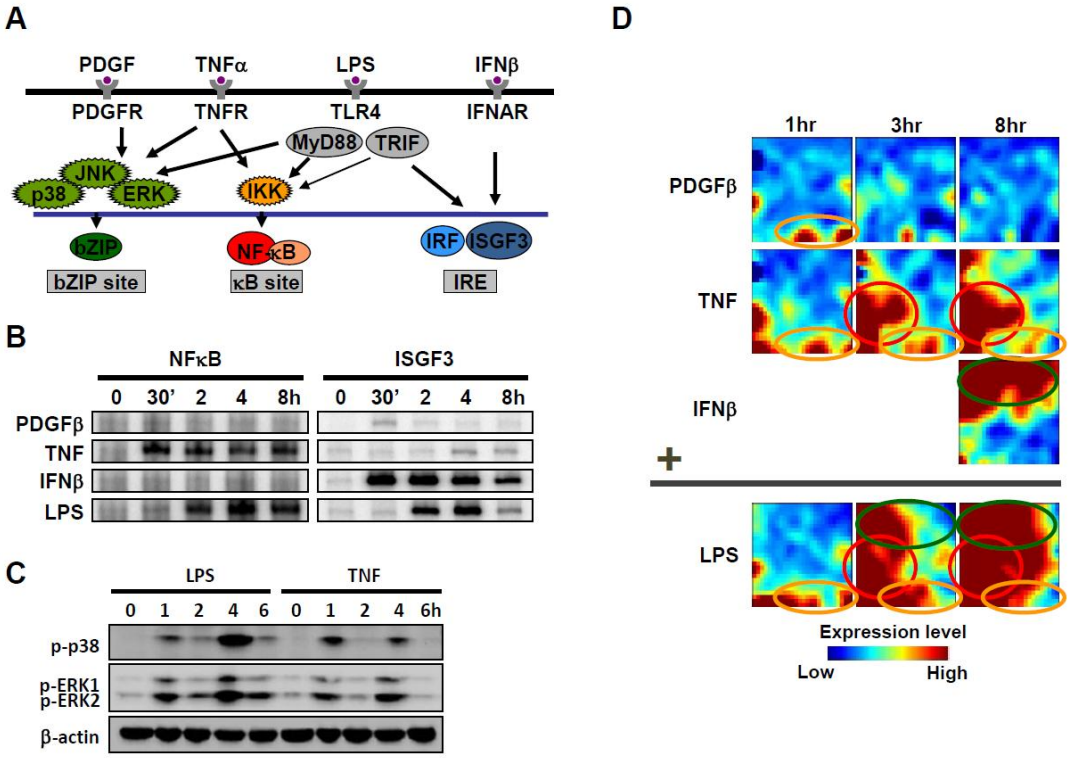


Figure 2.1 Dissect the LPS program by less complex cellular stimuli

(A) A schematic showing PDGF, TNF, LPS and IFNβ activated signaling pathways and downstream transcription factors and binding sites. (B) NFκB and ISGF3 DNA binding activities in response to PDGFβ, TNF, LPS and IFNβ at indicated times were revealed by EMSA with nuclear extracts from MEFs. (C) Phospho-p38 (Thr180/Tyr182) and phospho-ERK1/ERK2 (Thr202/Tyr204) were quantified by immunoblot with RIPA lysates prepared from MEFs at indicated times. (D) Microarray mRNA expression data from MEFs stimulated with PDGFβ (50 ng/ml), TNF (10 ng/ml), IFNβ (2500 units/ml) and LPS (0.1 ug/ml) at indicated times were analyzed by Gene Expression Dynamic Inspector that uses a self-organizing map (SOM).

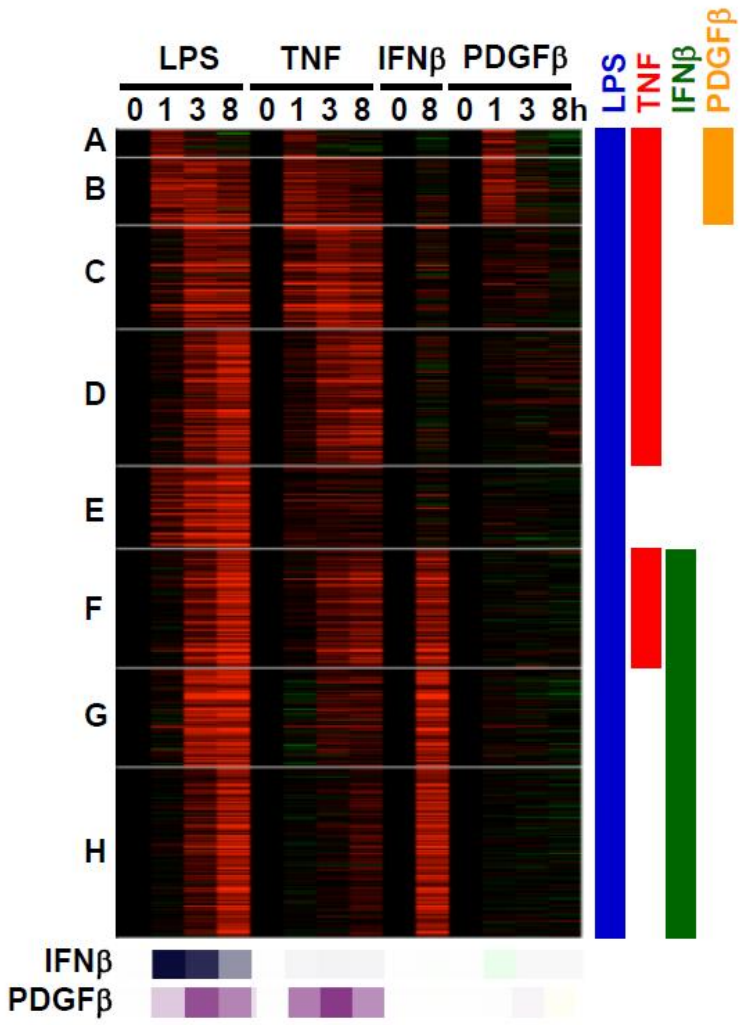


Figure 2.2 Three distinct groups of genes

Microarray mRNA expression data from MEFs stimulated with PDGFβ (50 ng/ml), TNF (10 ng/ml), IFNβ (2500 units/ml) and LPS (0.1 ug/ml) at indicated times were analyzed by K-means clustering. Expression fold change (log_2) was plotted as heatmaps, red represent gene induction and green represent gene repression. IFNβ (blue bar) and PDGFβ (purple bar) expression were presented at the bottom.

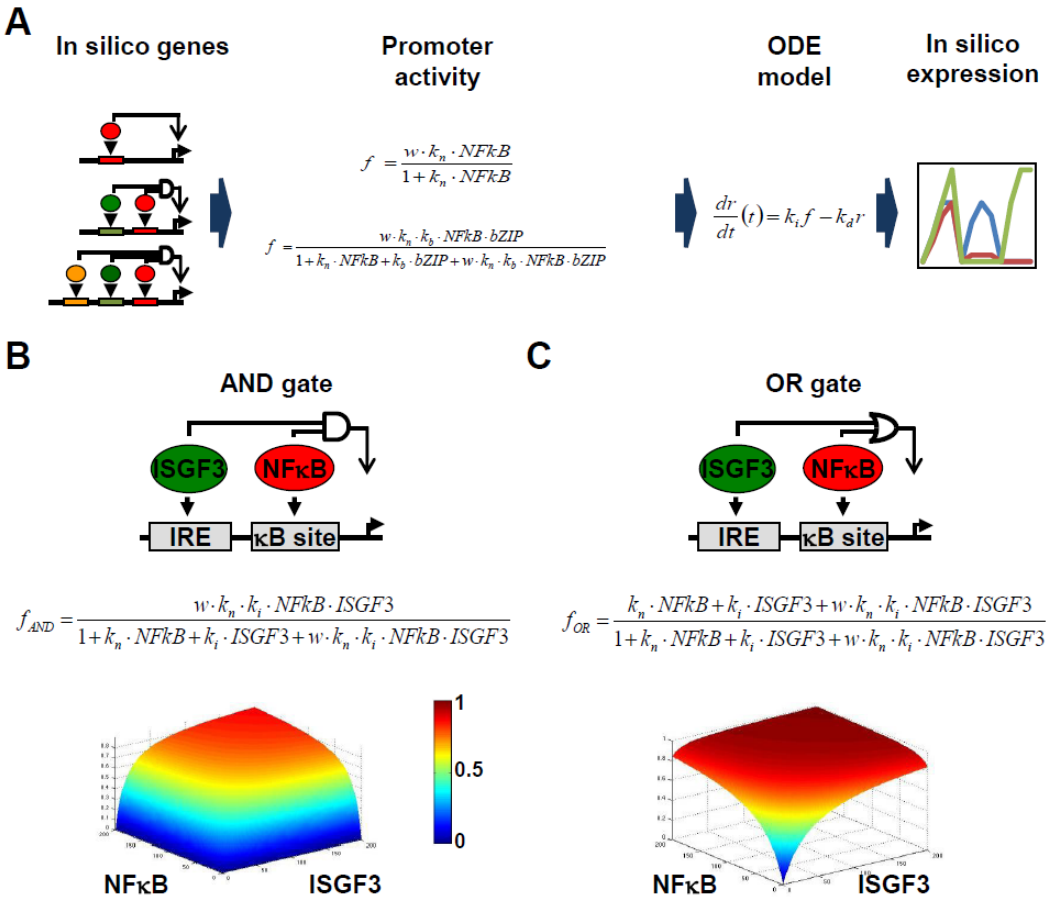
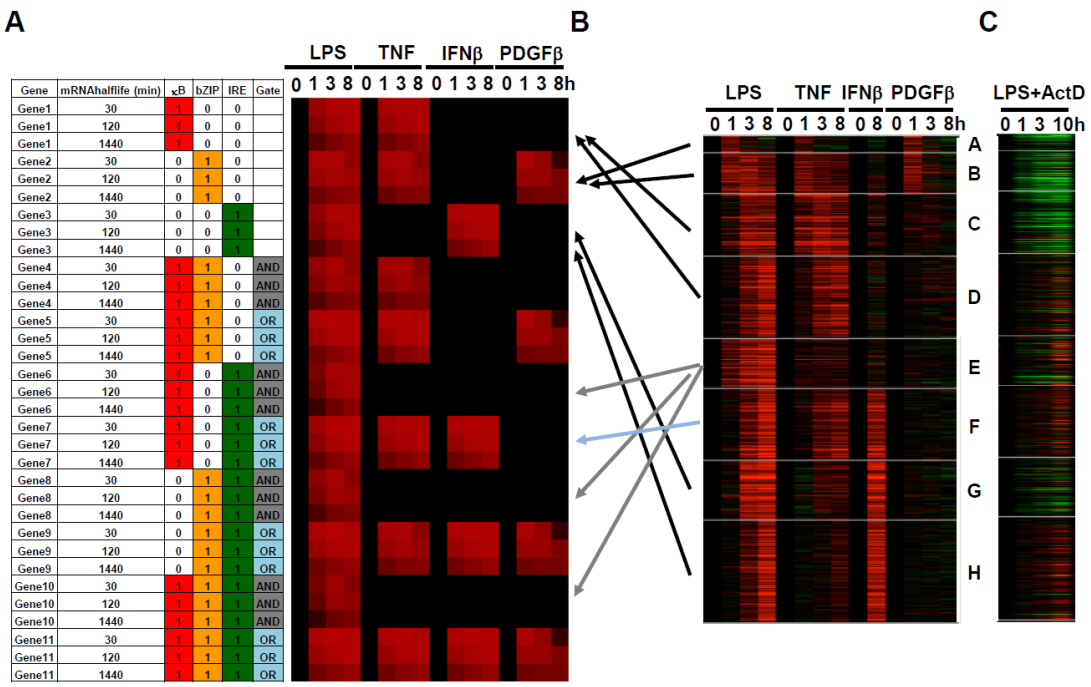


Figure 2.3 Construct mathematical models of in silico genes

(A) A schematic showing the construction of mathematical models for a set of stereotypical promoter architectures and the generation of in silico gene expression profiles. (B) A schematic and thermodynamic formulation of promoter activity f of an AND gate promoter that requires the activation of both ISGF3 and NFκB to induce gene expression. The mathematical solution of promoter activity (f) as a function of NFκB and ISGF3 concentrations. The color bars represent promoter activity (f) ranging from 0 (blue) to 1 (red). (C) A schematic, thermodynamic formulation and the mathematical solution of promoter activity f of an OR gate promoter that can be activated by either NFκB or ISGF3.



Gene	mRNAhalflife (min)	κB	bZIP	IRE	Gate
Gene1	30	1	0	0	
Gene1	120	1	0	0	
Gene1	1440	1	0	0	
Gene2	30	0	1	0	
Gene2	120	0	1	0	
Gene2	1440	0	1	0	
Gene3	30	0	0	1	
Gene3	120	0	0	1	
Gene3	1440	0	0	1	
Gene4	30	1	1	0	AND
Gene4	120	1	1	0	AND
Gene4	1440	1	1	0	AND
Gene5	30	1	1	0	OR
Gene5	120	1	1	0	OR
Gene5	1440	1	1	0	OR
Gene6	30	1	0	1	AND
Gene6	120	1	0	1	AND
Gene6	1440	1	0	1	AND
Gene7	30	1	0	1	OR
Gene7	120	1	0	1	OR
Gene7	1440	1	0	1	OR
Gene8	30	0	1	1	AND
Gene8	120	0	1	1	AND
Gene8	1440	0	1	1	AND
Gene9	30	0	1	1	OR
Gene9	120	0	1	1	OR
Gene9	1440	0	1	1	OR
Gene10	30	1	1	1	AND
Gene10	120	1	1	1	AND
Gene10	1440	1	1	1	AND
Gene11	30	1	1	1	OR
Gene11	120	1	1	1	OR
Gene11	1440	1	1	1	OR

Figure 2.4 Identify prevalent promoter architectures by correlating in silico and in vivo gene expression profiles

(A) A table showing 11 distinct promoter architectures with mRNA half-lives of 30, 120 and 1440 minutes. In silico mRNA expression profiles in response to LPS, TNF, IFNβ and PDGFβ generated from mathematical models of each gene were presented as heatmaps. (B) K-means clustering of microarray data from MEFs as shown in Figure 2 were matched with in silico expression profiles. (C) Microarray mRNA expression data from MEFs stimulated with LPS (0.1 ug/ml) for 3 hours and treated with actinomycin D (10 ug/ml) for 0, 3 and 10 hour were presented as heatmaps following the same order as the clusters in figure 2A.

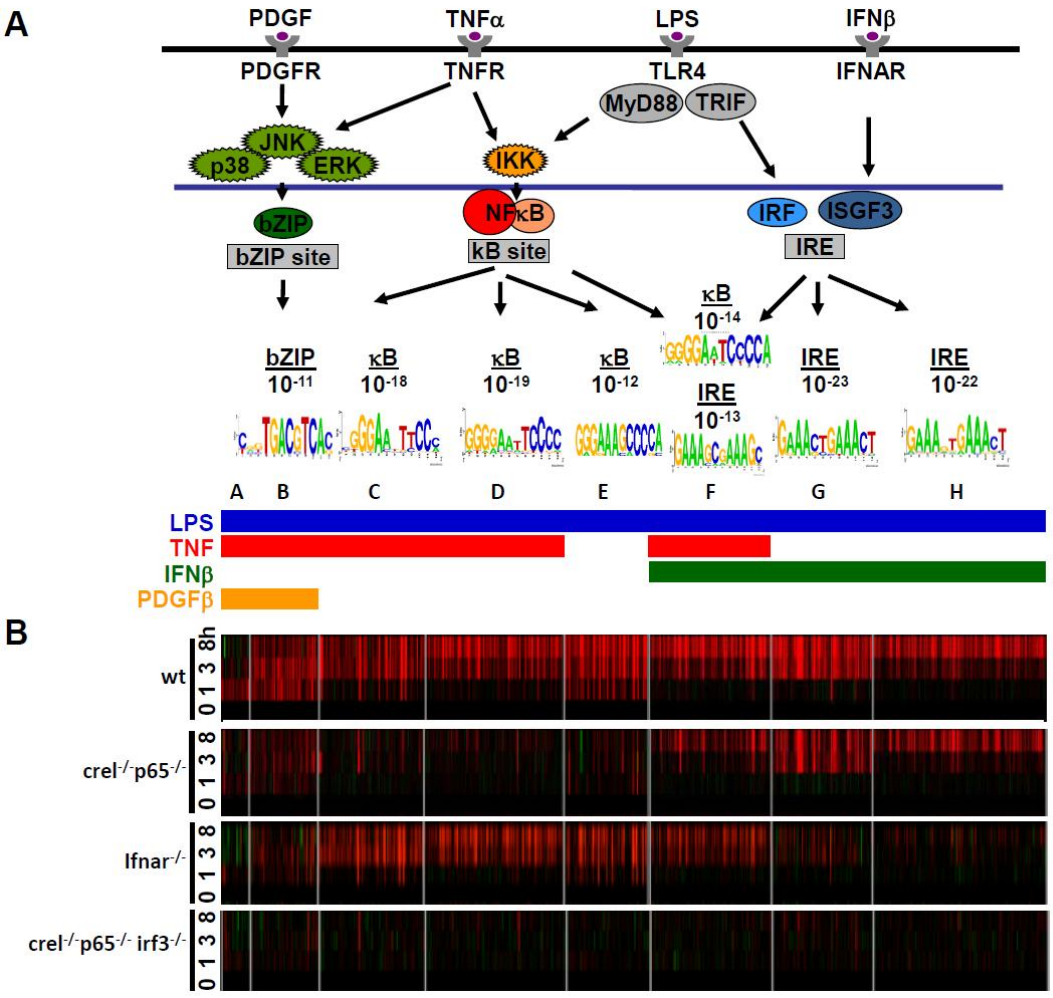


Figure 2.5 Bioinformatic and experimental analyses to validate model prediction

(A) A schematic of PDGF, TNF, LPS and IFNβ activated signaling pathway and transcription factors. Most highly enriched motifs identified *de novo* within -1.0 kb to 0.3 kb of transcriptional start sites are shown for each cluster with p-values to indicate statistical significance. Sequence logo of the enriched motif is presented below. (B) Microarray mRNA expression data from wild-type ("wt"), crel$^{-/-}$p65$^{-/-}$, ifnar$^{-/-}$, and crel$^{-/-}$ p65$^{-/-}$ irf3$^{-/-}$ MEFs stimulated with LPS (0.1 ug/ml) for 0, 1, 3 and 8 hour were presented as heatmaps following the same order as the clusters in figure 2A.

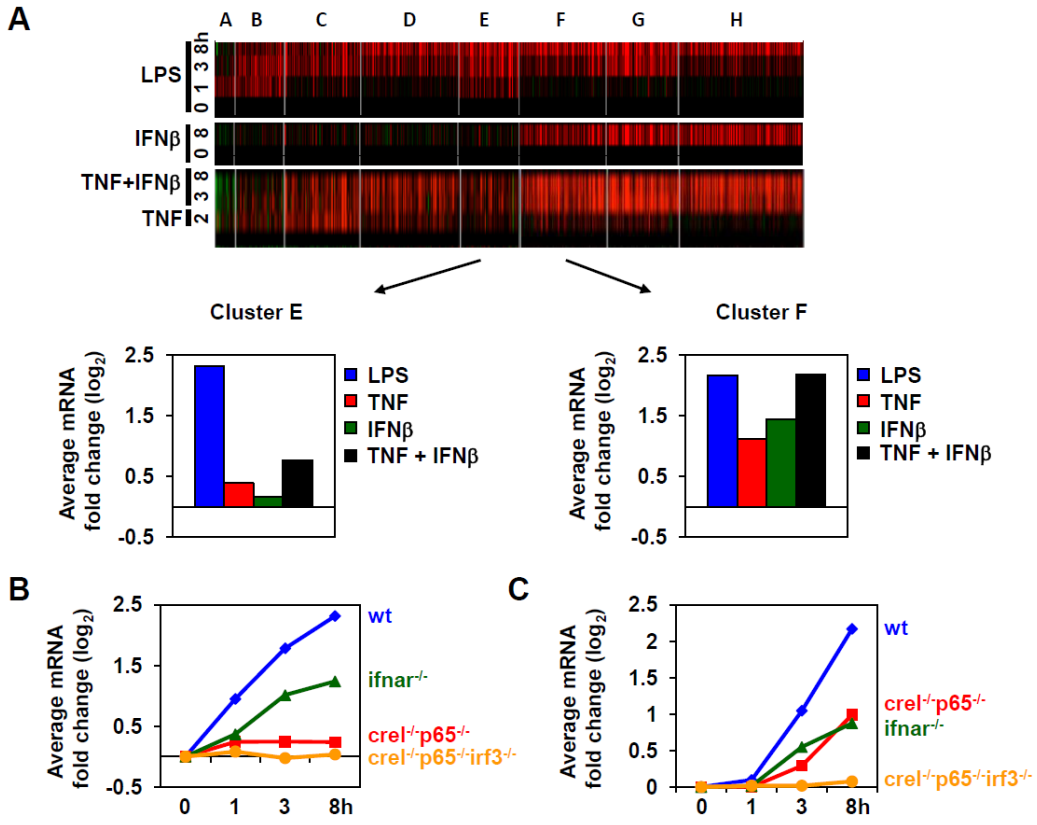


Figure 2.6 Finding experimental evidence for predicted OR and AND gate

(A) Microarray mRNA expression data from MEFs stimulated with LPS (0.1 ug/ml), IFNβ (2500 units/ml) or pre-stimulation with TNF (10 ng/ml) for 2 hour then stimulate with IFNβ (2500 units/ml) for 3 and 8 hour were presented as heatmaps following the same order as the clusters in figure 2A. Below, average mRNA expression fold change (log_2) of microarray data from cluster E and cluster F stimulated with LPS (blue), TNF (red), IFNβ (green) or TNF and IFNβ co-stimulation (black) for 8 hour were plotted as bar graphs. (B-C) Average mRNA expression fold change (log_2) of microarray data from cluster E (B) or cluster F (C) in wt (blue), crel$^{-/-}$p65$^{-/-}$ (red), ifnar$^{-/-}$ (green) and crel$^{-/-}$p65$^{-/-}$ irf3$^{-/-}$ (orange) MEFs stimulated with LPS (0,1 ug/ml) for 0, 1, 3 and 8 hour were plotted as line graphs.

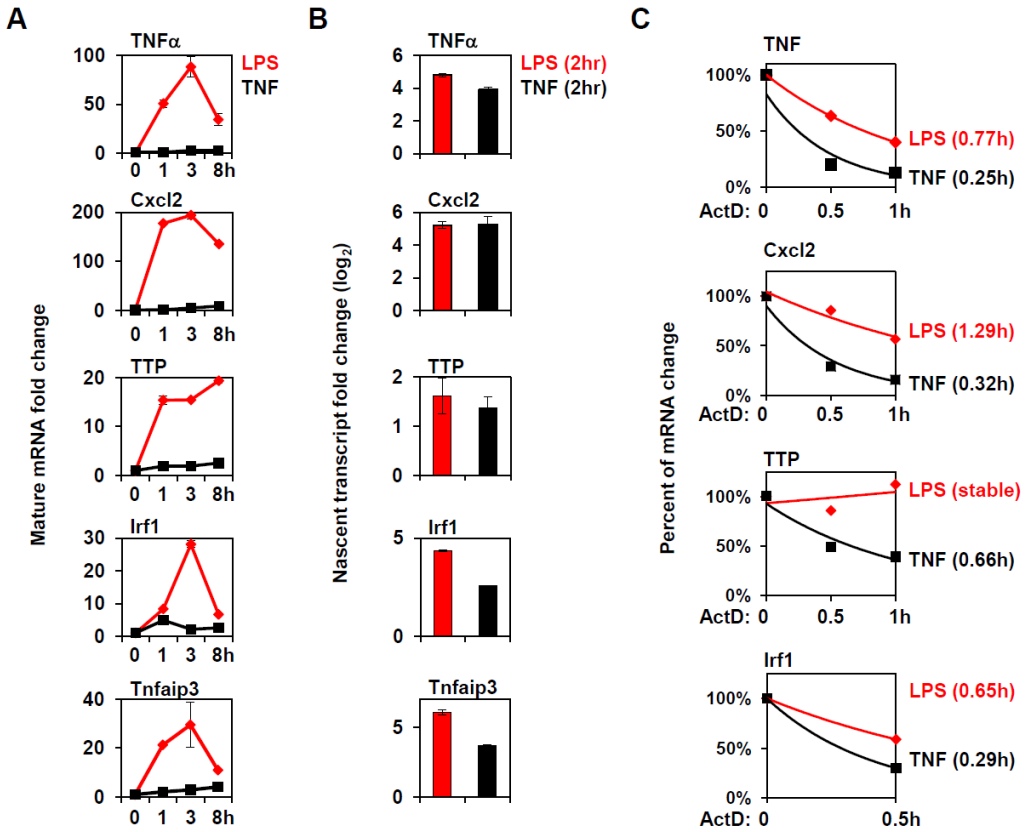


Figure 2.7 AND gate genes are controlled by regulated mRNA decay

(A) RT-qPCR of mature mRNA expression in bone marrow derived macrophages (BMDMs) stimulated with LPS (0.1 ug/ml) or TNF (10 ng/ml) for 1, 3 and 8 hour. Data was normalized to the expression of Gapdh and then plotted as expression fold change relative to un-stimulated cells. (B) Nascent transcript levels were measured by RT-qPCR of chromatin associated RNA in BMDMs stimulated with LPS (0.1 ug/ml) or TNF (10 ng/ml) for 2 hour. Data were represented as fold induction (log_2) relative to no-stimulated cells. (C) Measurement of mRNA half-lives were performed by RT-qPCR of mRNA expression in BMDMs stimulated with LPS (0.1 ug/ml) or TNF (10 ng/ml) for 30 minutes, then treat the cells with actinomycin D for 0, 0.5 and 1 hour. Data for LPS (red) or TNF (black) were presented as percent of mRNA decrease relative to without actinomycin D treatment. mRNA half-lives calculated from exponential regression analyses were shown in parenthesis.

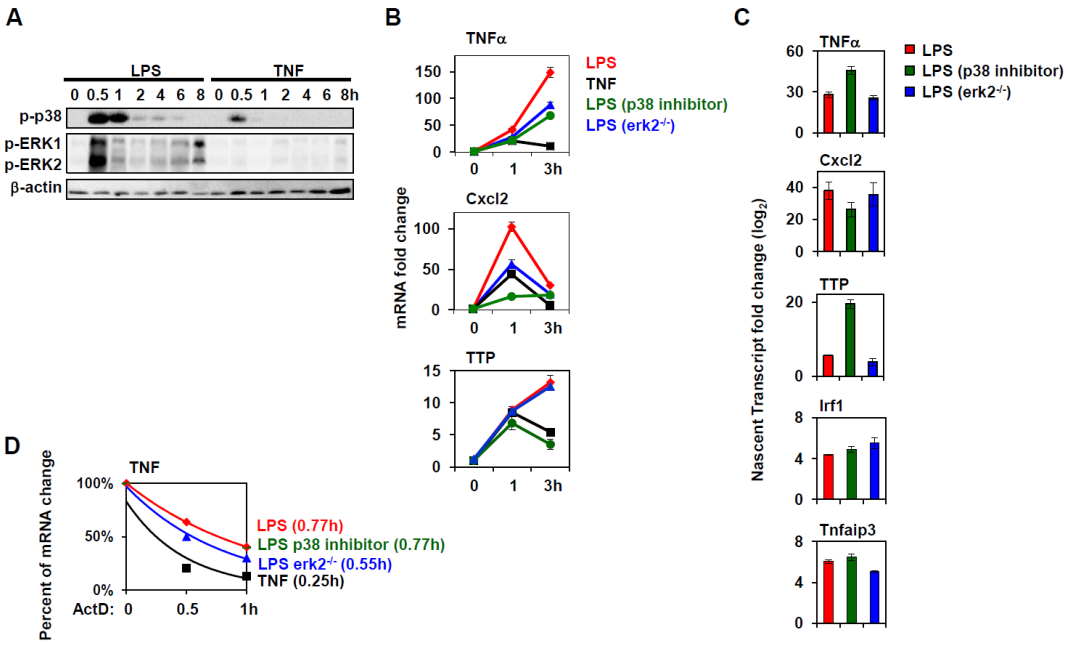


Figure 2.8 LPS specific induction of mRNA stabilization pathway

(A) Phospho-p38 (Thr180/Tyr182) and phospho-ERK1/ERK2 (Thr202/Tyr204) were quantified by immunoblot with RIPA lysates prepared from BMDMs at indicated times. (B) RT-qPCR of mature mRNA expression in wt or erk2$^{-/-}$ bone marrow derived macrophages (BMDMs) stimulated with LPS (0.1 ug/ml), TNF (10 ng/ml) or pre-treated with p38 inhibitor (SB202190) (2.5 uM) for 30 minutes then stimulated with LPS for 1, 3 and 8 hour. Data was normalized to the expression of Gapdh and then plotted as expression fold change relative to un-stimulated cells. (C) Nascent transcript levels were measured by RT-qPCR of chromatin associated RNA in wt or erk2$^{-/-}$ BMDMs stimulated with LPS (0.1 ug/ml), TNF (10 ng/ml) or pre-treated with p38 inhibitor (SB202190) (2.5 uM) for 30 minutes then stimulated with LPS for 2 hour. Data were represented as fold induction (log_2) relative to no-stimulated cells.

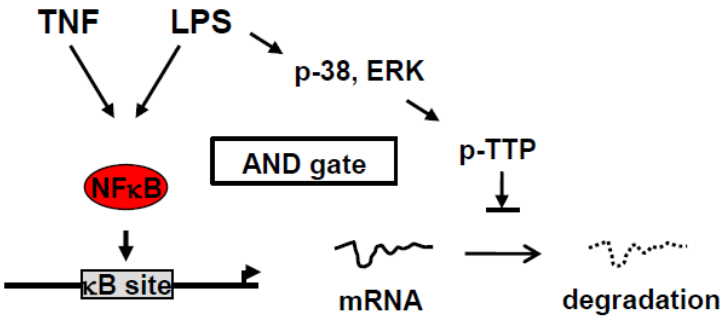


Figure 2.9 Synergy between transcription and regulated mRNA decay

A schematic showing our model of a synergistic AND gate between transcriptional regulation from NFκB and LPS specific activation of the mRNA stabilization pathway, which is activated through LPS induced phosphorylation of p38 and ERK and the subsequent phosphorylation of the mRNA de-stablization protein TTP.

Chapter 3 : The specificity of innate immune responses is enforced by NFκB p50 repressing interferon-regulatory elements (IREs)

ABSTRACT

Specific binding of transcription factors to cognate sequence elements is thought to be critical for the generation of specific gene expression programs. Nuclear Factor κB (NFκB) and the interferon regulatory factors (IRFs) are known to bind the κB-site and the interferon regulatory element (IRE), respectively, and are activated in macrophages following exposure to pathogen. However, how these factors produce pathogen-specific inflammatory and immune responses remains poorly understood. Combining top-down and bottom-up systems biology approaches, we identify the NFκB p50 homodimer as a regulator of IRF responses. First, unbiased genome-wide expression analysis revealed that p50 represses a subset of interferon-inducible genes via a newly defined subclass of guanine-rich IRE (G-IRE) sequences; this finding was substantiated by biochemical and structural analyses. Second, mathematical modeling predicted that p50 homodimer may enforce stimulus-specificity of composite promoters; indeed, the expression of the antiviral regulator IFNβ was found to be stimulus-restricted by p50 homodimer binding to the G-IRE-containing enhancer to suppress cytotoxic IFN signaling. Specifically, p50 deletion resulted in inappropriate production of IFNβ in response to bacterial DNA exposure sensed by TLR9. The novel role for NFκB p50 in enforcing the specificity of cellular response to pathogens by binding a newly defined subset of IRE sequences alters our understanding of how NFκB and IRF signaling systems cooperate in regulating anti-microbial immunity.

INTRODUCTION

The cellular innate immune response is triggered by three families of extra-cellular and intra-cellular receptors for a broad range of molecular patterns (Takeuchi and Akira, 2010). The Toll-like receptors (TLRs), the RIG-I-like receptors (RLRs) and the NOD-like receptors (NLRs) show specificity in the recognition of pathogen associated molecular patterns (PAMPs) and metabolic products, termed danger-associated molecular patterns (DAMPs). Although many cell types are capable of innate immune responses, macrophages are professional, tissue resident initiators, coordinators, and effectors of the innate immune response.

Pathogen recognition by macrophages elicits gene expression programs typically consisting of hundreds of genes (Doyle et al., 2002; Nau et al., 2002; Ogawa et al., 2005), that may be broadly classified as mediating cellular anti-viral functions and systemic immune activation via inflammation (Foster et al., 2007). Because both responses are also potentially detrimental to the organism (Decker et al., 2005; Marshak-Rothstein, 2006; Nathan, 2002), cells are thought to produce pathogen-specific responses, ensuring that unnecessary gene products remain turned off. Dozens of transcription factors have been implicated in providing for fine-tuned innate immune gene expression programs (Ramsey et al., 2008; Roach et al., 2007). However, what the critical regulators are in producing pathogen- or stimulus-specific gene expression remains an open question that is of relevance to an understanding of anti-microbial immune responses, chronic

inflammatory disease and the development of relevant therapeutics, including adjuvants for innate and adaptive immune responses.

Two transcriptional regulators that play central roles in coordinating the cellular innate immune response expression programs are the nuclear factor κB (NFκB) and the interferon regulatory factors (IRF) (Akira and Takeda, 2004; Akira et al., 2006; Amit et al., 2009; Doyle et al., 2002; Medzhitov, 2008; Oganesyan et al., 2006) (Figure 3.1). Interestingly, stimulus-responsive activation of NFκB and IRFs are determined by signaling adaptors that selectively interact with the intracellular domains of TLRs (Kawai and Akira). MyD88 mediates the activation of NFκB and is associated with the majority of the TLRs, including TLR9, the bacterial CpG-rich DNA sensor, and TLR4, the sensor for the gram-negative cell wall component lipo-polysaccharide. TRIF functions as the primary signaling adaptor for IRF activation and is associated with many TLRs, including TLR4, but not TLR9. Hence, each TLR produces a characteristic combination of transcription factor activities, including NFκB and IRF.

Both NFκB and IRFs constitute families of transcription factors that share highly conserved DNA binding domains. In macrophages, IRF3 is activated through site-specific phosphorylation by TBK1 (Sharma et al., 2003), an important effector kinase of the TRIF pathway. IRF3-driven type I interferon production results in an autocrine loop that activates a second IRF family member, ISGF3, whose DNA binding component is IRF9. IRF3 and ISGF3 appear to have largely overlapping DNA binding specificities for

the IRE consensus (AANNGAAA) (Fujii et al., 1999). Within the NFκB family, the key transcriptional effectors are the activation domain-bearing RelA and cRel proteins, which as dimers with the nfkb1 gene product p50, are responsible for κB-driven gene activation. Overlapping DNA binding specificities for the broad κB consensus GGRNNN(N)YCC (Hoffmann et al., 2006) underlie the fact that there have only been isolated reports of specific cRel or RelA target genes (Hoffmann et al., 2003; Sanjabi et al., 2000).

However, another major NFκB family member is the p50:p50 homodimer, a presumptive transcriptional repressor of κB-sites by virtue of its close sequence and structural homology with other NFκB family members and the lack of a transcriptional activation domain (Kang et al., 1992; Ten et al., 1992). Indeed, p50:p50 homodimer was shown to function as a competitive repressor of κB-driven transcription in transient transfection and *in vitro* studies (Ledebur and Parks, 1995; Plaksin et al., 1993), and may repress the expression of the *tnf* gene (Bohuslav et al., 1998; Sha et al., 1995; Udalova et al., 2000). However, the binding specificity and physiological functions of this putative transcriptional repressor remain surprisingly uncharacterized. Interestingly, while p50's role as a dimerization partner for RelA and cRel can be compensated for by the nfkb2 gene product p52 (Hoffmann et al., 2003), p52:p52 homodimers have not been observed, indicating that p50:p50 may play unique functions as a repressor within the NFκB transcription factor family. Within the innate immune response, where inappropriate gene expression is potentially detrimental for cellular or organismal health, the role of transcriptional repressors is of pertinent interest. To characterize the physiological role

of the p50:p50 homodimer in the innate immune response, we first undertook unbiased genome wide studies to identify functional targets, and then characterized the molecular mechanism both at the level of biophysical DNA binding specificity and within the context of the gene regulatory circuitry of composite enhancer elements.

RESULTS

NFκB p50 represses IFN response genes

To characterize the functional role of the NFκB p50:p50 homodimer in the immune response in macrophages, we first determined its abundance by quantitative Western blotting. Using recombinant protein standards we estimate 30,000 of these transcriptional-activation domain-lacking homodimers in the nuclei (i.e. 200nM) of naïve or resting bone marrow derived macrophages (Figure 3.2). To identify functional targets of the p50 homodimer during the cellular response to pathogens in an unbiased manner, we profiled gene expression induced by bacterial lipopolysaccharide (LPS) in *nfkb1*$^{-/-}$ ("p50ko") cells using microarray analysis. We used K-means clustering to analyze relevant microarray datasets generated with murine embryo fibroblasts (MEFs) and bioinformatically identified shared sequences in the promoters of co-regulated genes (Figure 3.3). To our surprise we found that p50ko cells showed hyper-expression of genes not known to be NFκB-regulated (Figure 3.3, clusters I, J and K). We found that these genes to be interferon-inducible (last column) and to contain IREs in their promoter sequences. Similar studies with bone-marrow-derived macrophages (BMDMs) also revealed a p50-mediated repressive effect on sets of IRE-containing genes in response to

LPS (Figure 3.4, clusters B, C, E) or in response to moderate interferon-β stimulation (Appendix Figure 3.1, clusters B and C). Interestingly, p50's known function as a binding partner to transcriptional activator RelA manifested itself in only a small reduction of some NFκB-regulated genes (Figure 3.3, clusters C and D, Figure 3.4, clusters D and F), indicating that it is redundant with the nfkb2 protein p52 or RelA homodimers, or that p50:p50 repressive functions mask p50:RelA activation functions (Hoffmann et al., 2003).

The *nfkb1* gene gives rise to two gene products, the mature p50 protein and its precursor, p105, which may sequester activating NFκB subunits (Liou et al., 1992) and participate in ERK signaling (Beinke et al., 2003; Waterfield et al., 2003). Examining LPS induced gene expression profiles in wt and *p105*$^{-/-}$ BMDMs, which lack the p105 precursor but express p50 (Chang et al., 2009), we did not observe elevated expression of interferon inducible genes (Appendix Figure 3.2). Our results suggest that the mature p50 protein participates in the regulation of interferon response genes, either directly or indirectly. To characterize the proteins that bind IRE targets, we performed electrophoretic mobility shift assays (EMSA) with a widely used IRE probe (Koenig Merediz et al., 2000). Time course studies confirmed that LPS activates the transcriptional activators NFκB (p65:p50), IRF3 and subsequently ISGF3 (IRF9 associated with Stat1/Stat2), while the κB binding activity of the transcriptional repressor p50 homodimer is constitutive in resting cells and is increased at late time points (Figure 3.5 and Appendix Figure 3.3) due to NFκB (p65:p50)-induced transcriptional activation

of the nfkb1 gene (Cogswell et al., 1993). To our surprise, we observed the formation of an unexpected complex ("complex X") on the IRE probe (Figure 3.5). Complex X did not only have the same mobility as the p50 homodimer on the κB probe, but was found to be absent in p50-deficient BMDMs (Figure 3.6). Antibody supershifts showed that this complex contained p50 but not p65 (Figure 3.7). Cold probe competition assays show that the IRE and κB probes cross-compete in binding the p50 homodimer (Figure 3.8), and titration of recombinant p50 on IRE and κB probes indicated that the p50 homodimer has similar affinity for IRE and κB site sequences (Figure 3.9). Taken together, our results suggest that the constitutively present p50 homodimer may repress IFN–inducible genes by directly binding to IRE sites.

NFκB p50 regulates and binds "G-IREs"

Examining the promoter sequences of well-known IRF target genes, we found that genes that showed marked hyperexpression in p50ko cells tend to contain guanine-rich sequences within or next to their IRE (Figure 3.10A, red), whereas others that showed little misexpression in p50ko cells do not (Figure 3.10A, black). To evaluate the role of the G-rich sequence statistically, we generated replicate microarray datasets from LPS-induced BMDMs. Focusing on interferon-inducible genes whose activation by LPS was IFNAR-dependent, we found a continuum of hyper-expression phenotypes in LPS-induced p50ko cells (Figure 3.10B), consistent with a range of relative p50 and IRF DNA binding affinities or numbers of binding sites. Marking genes whose regulatory regions contain G-rich IREs ("G-IREs") showed that G-IREs were over-represented in

hyperactivated genes (Figure 3.10B). Exceptions to this correlation were also apparent and may be due to transcriptional saturation effects, misregulation of the ERK pathway (Beinke et al., 2003; Waterfield et al., 2003) or interferon expression, or the likely possibility that the G-rich classification may in some cases consider IRE-like sequences that are not functionally relevant. However, statistical evaluation indicated (with a *p*-value of $<10^4$) that G-IRE-containing interferon-inducible genes were more likely to be misregulated in the p50ko cells than non-G-rich-containing counterparts (Figure 3.10C). Furthermore, endogenous p50 homodimer formed complexes with DNA probes designed to contain the IRE-containing regulatory sequences derived from several p50-repressed genes, but not when analogous sequences from p50-independent genes were used (Figure 3.11). Chromatin immunoprecipitation (ChIP) assays revealed p50 recruitment *in vivo* to those endogenous IRF target genes that are hyper-expressed in p50-deficient cells, but less to those IRF target genes that are largely independent of p50 (Figure 3.12A). Control experiments confirmed that neither class of IRF target genes recruits the NFκB p65 protein (Figure 3.12B), whereas the ISGF3 transcription factor was found recruited to both classes of genes (Appendix Figure 3.4).

Based on NFκB crystal structures bound to near palindromic κB sites (Chen et al., 1998a; Ghosh et al., 1995), we modeled a p50-homodimer on the GBP-1 G-IRE/IRE sequences (Figure 3.13A); one monomer is likely to make base-specific contacts with the G-IRE sequence which conforms to the one half-site κB site consensus (5'-GGRN-3') (Figure 3.13B), whereas the second monomer does not make base-specific contacts with the second IRE as it deviates significantly from the κB consensus (Figure 3.13C).

However, the central tyrosine clamp may contribute affinity to both monomers, as it is largely sequence independent. Interestingly, the hinge region between the DNA binding and dimerization domains within the Rel homology region of the second monomer assumes a different angle in the two monomers. Such structural considerations suggest that a single half-site containing 3 or 4 guanines provides sufficient number of contacts for high affinity binding (Figure 3.13B), whereas the second monomer may accommodate suboptimal sequences by avoiding disallowed contacts (Figure 3.13C) due to rotational flexibility (about 18°) between the DNA-binding and dimerization domains of the p50 protein (Figure 3.13A). Indeed, mutating the binding site confirmed that the -3 position G is critical for purified recombinant p50 homodimer binding, whereas the central adenines are not (Figure 3.14). Our analysis suggests a model of p50-IRE binding wherein one monomer interacting with a G-IRE provides sufficient affinity and the second monomer has rotational flexibility to accommodate a non-consensus second half-site sequence.

Functional consequences of p50:p50-G-IRE interactions

We studied the immediate functional consequences of p50 homodimer binding to G-IRE sites. Titrating recombinant p50 into *in vitro* binding assays revealed that both ISGF3 and IRF3 shifts are competed (Figure 3.15A); neither p65 titration nor "G-rich"-lacking DNA sequences showed this competition effect (Figure 3.15B), and no higher order complexes containing both p50 and IRF proteins were ever observed. Transient transfection studies revealed that expression of only a DNA-binding competent p50 reduced IRF-mediated activation of a G-IRE-driven reporter gene specifically (Figure

3.16), but not a non-G-IRE reporter gene (Appendix Figure 3.5). Our results suggest that p50:p50 may inhibit promoter activity by simply competing with IRF3 and ISGF3 in binding G-IREs, although an active repression mechanism via recruitment of co-repressors cannot be ruled out.

To explore the functional consequences of the competition model for promoter activity, we used a thermodynamic model (Bintu et al., 2005b) to calculate fractional promoter activity as a function of p50 homodimer and IRF activator concentrations (Figure 3.17A). The model recapitulated our experimental data that p50-repressed IRE containing genes are hyper-expressed in p50ko cells, with the relative effect being more substantial at low IRF concentrations (Figure 3.17B) as they may occur in the early phases of a time course. Indeed, a simple parameter sensitivity analysis confirms that hyperactivation in p50ko cells is not only a function of the affinity of p50 homodimer to the IRE but is also more pronounced at low IRF concentrations (Figure 3.17C). Interestingly, G-IRE-containing promoters were more likely to show elevated expression in unstimulated conditions in p50ko cells (Appendix Figure 3.6 nad 3.2), a phenomenon that relates to the previous gene expression phenotype, presumably due to low levels of basal IRF3 activity,

Using the mathematical modeling approach we wished to explore the functional consequences of p50:p50 regulating G-IREs on physiologically relevant gene regulatory circuits. Many innate immune response genes contain clusters of IRE and κB-sites that are thought to function cooperatively to allow for stimulus-specific gene control (Fan and

Maniatis, 1989; Mark Ptashne, 2002; Neish et al., 1995). Such functional cooperativity may be described with a thermodynamic model for so-called AND-gate promoters (Bintu et al., 2005a; Bintu et al., 2005b). We derived an expression for the promoter activity of a gene whose duplicate IREs and a κB-site synergize via cooperative IRF and NFκB interactions and where only the κB site is subject to p50:p50 repression (Figure 3.18A). As expected, the formulation shows that both NFκB and IRF must be activated to induce the promoter; however, the promoter activity heatmap shows that even low amounts of IRF activity may be sufficient to render the promoter responsive to NFκB activation. Considering that IRF3 and ISGF3 can be found constitutively in the nucleus of even unstimulated cells (for example Appendix Figure 3.7), this may be of physiological relevance, as PAMPs may be distinguished as those activating both NFκB and IRF, and those activating only NFκB (Akira and Takeda, 2004; Akira et al., 2006). For example, macrophages exposed to bacteria-derived CpG-containing DNA activate NFκB whereas exposure to endotoxin LPS activate both NFκB and IRF factors. Thus non-zero basal activity of IRF3 or ISGF3 may result in inappropriate promoter induction in response to stimuli such as CpG that only activate NFκB. The modeling illustrates that AND-gates may be rendered surprisingly leaky when inducible transcription factors show significant basal activity.

We constructed an alternate thermodynamic model to explore the functional consequence of the p50 homodimer repressing a G-IRE within this AND-gate promoter architecture (Figure 3.18B). The analogous promoter activity heatmap shows that the

CpG-responsive expression is effectively abolished, but LPS responsiveness is preserved. The result is a sharpened stimulus-specificity of the promoter, suggesting that the p50 homodimer may have gate-keeping functions on G-IRE-containing composite promoters. We examined whether this conclusion is dependent on specific parameter values within these models of hypothetical response genes. A multi-dimensional parameter sensitivity analysis of the four relevant dissociation constants indicated that LPS vs. CpG stimulus-specificity is enhanced by p50:p50 homodimers repressing the G-IRE in a wide variety of conditions and never diminished. Such promoter "gating" to restrict gene expression to specific stimuli is independent of p50:p50 binding to the κB site (see low $K_{p50\text{-}IRE}$) and is strongest on promoters where IRFs have moderate but p50:p50 have strong binding affinity (Figure 3.19). We note that G-IREs have high affinity for p50:p50 (Figure 3.9) but may be suboptimal for IRFs.

NFκB p50:p50 restricts anti-viral responses to specific stimuli

To test the prediction that the p50 homodimer may enforce stimulus-specificity of AND-gate promoters, we examined one of the most well studied and physiologically important composite promoters, the one controlling the expression of interferon-β (IFNβ). Synergistic function by NFκB and IRF transcription factors via coordinated DNA binding sites (so called "PRD" elements) is known to be critical for IFNβ gene control (Escalante et al., 2007). Interestingly, p50 was reported to bind the IFNβ enhancer in unstimulated cells (Parekh and Maniatis, 1999), but the ChIP-assay does not allow for sufficient resolution to determine its precise location(s). Interestingly, within

the regulatory sequence not only the κB site in PRDII but the IREs in the PRDI and PRDIII regions contain G-rich sequences (Figure 3.20). Indeed, a DNA binding probe that encompasses all PRD elements was found to complex not only one but two p50 homodimers (lanes 1-3). Dissecting their location by mutating the three G-rich stretches we confirmed that the κB site recruits one p50 homodimer (lanes 4-9), but that the central triple G element within the IREs recruits a second p50 homodimer (lanes 10-18). Though this triple G element is not required for IRF binding, it is evolutionarily conserved (Panne et al., 2007). To explore the functional consequence of these binding events, we constructed a mathematical model of IFNβ mRNA expression, in which the thermodynamic formulation of promoter activity *f* (Figure 3.21A) is embedded in a differential equation describing mRNA production and decay. Simulations show that this model predicts a significant differential responsiveness of the IFNβ gene to CpG and LPS stimulation in wild type but not p50ko cells (Figure 3.21B). Indeed, our experimental analysis reveals that, whereas wild type cells do not allow for IFNβ expression in response to CpG, p50ko cells show substantial misexpression (Figure 3.22).

As IFNβ coordinates a large antiviral gene expression program via the activation of ISGF3, we examined how stimulus-specific gating of the IFNβ enhancer by the p50 homodimer would affect the downstream physiology (Figure 3.23). We characterized LPS and CpG induced ISGF3 activation profiles and found inappropriate induction of ISGF3 activities in response to CpG in the p50ko macrophages (Figure 3.24A), while IRF3 was not activated (Appendix Figure 3.6). Furthermore, CpG-induced the expression

of known anti-viral genes (Figure3.24B), that are in fact part of a large group of IFN-response genes revealed by microarray studies (Figure 3.25). We established a viral infection assay in which IFNβ priming of macrophages inhibited the infectivity of a GFP-expressing cytomegalovirus (CMV) (Mathys et al., 2003), whereas IFN-signaling defective macrophages were more susceptible (Appendix Figure 3.6). Exposing macrophages to LPS rendered them more resistant to CMV infection whereas CpG exposure did not (Figure 3.26A and 3.26B), reflecting the stimulus-specific expression of IFNβ. Such stimulus-specificity in mounting viral resistance was severely compromised in p50ko macrophages, as CpG exposure also resulted in increased anti-viral resistance in p50-deficient macrophages. Our results may relate to resistance of *nfkb1*$^{-/-}$ mice to encephalomyocarditis virus (EMCV) (Sha et al., 1995) and of an *nfkb1*$^{-/-}$ immortalized fibroblast cell line to influenza (Wei et al., 2006). Given IFN's cytostatic or even cytotoxic effects, stimulus-restricted expression of IFNβ is important for the health of cells and organisms. Whereas CpG promotes cell proliferation in wild type macrophages, we found that in p50ko macrophages it causes cytotoxicity as a detrimental consequence of a hyperactivated antiviral immune program (Figure 3.27).

DISCUSSION

Our findings revealed unexpected cross-regulation between the two primary transcription factors that coordinate innate immune responses. Whereas the NFκB and IRF activators bind their respective cognate sites (κB and IRE, respectively), we report

that the NFκB p50:p50 repressor binds and regulates a subset of IREs, termed G-IREs, whose sequence contains or is in direct vicinity to a three-guanine sequence. We show that as an abundant constitutive component of the nucleus, p50 homodimers have significant thresholding functions for IRF-responsive gene expression. Exemplified by the combinatorial IFNβ enhancer, our work emphasizes the importance of homeostatic repressors in restricting gene expression to specific stimuli by binding to sites that overlap with those of stimulus-inducible factors.

By combining unbiased gene expression phenotyping studies of *nfkb1*$^{-/-}$ cells with biochemical and molecular biological studies we provided *in vitro* and *in vivo* binding evidence that p50 directly regulates G-IRE containing promoters. However, indirect mechanisms such as hyper-expression of autocrine type-I interferon, or regulation of ERK by the p105 precursor p50 may contribute to the gene expression phenotype, and may provide an explanation for the hyperactivation of genes whose proximal regulatory region (-1kb to +0.3kb) we did not find a G-rich sequence flanking an IRE. Given that IFN-responsive ISGF3 was not hyper-activated in p50ko cells in response to LPS (Figure 3.24A), that the hyperactivation was observed at 1hr (Figure 3.3 and 3.4), i.e. prior to ISGF3 activation, and that many G-IRE genes were hyper-responsive to ectopic IFN stimulation (Appendix Figure 3.1) corroborates our conclusion that p50 regulates G-IRE-containing promoters directly.

The sequence specific interaction of transcription factors with their specific DNA binding sites is an organizing principle for understanding the logic of gene regulatory

circuits. Sequence specificity has traditionally been determined using low throughput biochemical assays, or PCR-mediated DNA selection schemes that identify the highest affinity DNA sequences (Kunsch et al., 1992). However, recent unbiased ChIP-chip or ChIP-seq experiments have revealed association of transcription factors with DNA that do not contain known cognate sequences, although indirect binding cannot be ruled out (Martone et al., 2003; Schreiber et al., 2006). High throughput biochemical affinity measurements in miniaturized assay systems have however, revealed that some transcription factors have much broader binding specificity than anticipated or may have more than one mode of specific DNA binding (Badis et al., 2009; Maerkl and Quake, 2007; Ragoussis et al., 2006; Warren et al., 2006). These observations set the stage for overlap in the sequence space associated with different transcription factor families. Competitive binding between specific IRF and NFκB family members is an example of this scenario with important gene regulatory consequences. Our study suggests that DNA elements have evolved to recruit two distinct DNA binding proteins to achieve specific regulation of expression; such hybrid elements may not be uniquely assigned as a response element of a single signaling pathway. In the case of G-IREs, we classify these sequences as a subclass of IREs, rather than of κB sites, because the G-IRE conforms to the direct repeat character of IREs unlike the palindromic κB site, and IRFs are the cognate transcriptional activators, whereas the NFκB family member appears to function as a competitive repressor.

Delineating the specificity of transcriptional regulators remains challenging. High-throughput cell-free biochemical approaches characterize the capacity of DNA

binding domains to bind sequences, but may be hampered by non-physiological reaction conditions including the quality of the recombinant protein (Badis et al., 2009; Maerkl and Quake, 2007; Ragoussis et al., 2006; Warren et al., 2006). We used high-throughput gene expression measurements in genetically altered cells to identify candidate interactions that are functional. Biochemical studies confirmed the capacity of the p50:p50 homodimer to bind G-IREs and to repress IRF-driven transcription. Biophysical considerations based on related X-ray structures rationalize the biophysical basis for p50:p50 to bind a triple G-motif without requiring a second palindromic halfsite; namely, high affinity interactions by one monomer and a flexible hinge between the DNA binding and dimerization domains within the other. In contrast, the ubiquitous NFκB activator p50:p65 dimer does not form a complex on G-IRE probes (Figure 3.5). Previous crystallographic studies revealed fewer hydrogen bonds by p65 than p50 with a single half-site suggesting lower affinity binding (Chen et al., 1998b), and future studies ought to address whether p65's hinge may not be as flexible as p50's. However, which IRF family member the p50-homodimer competes with is less unambiguous. We observed a stronger gene expression phenotype in resting p50ko cells and at early time points following LPS exposure (Figure 3.3, 3.4 and Appendix Figure 3.2), correlating with the activation profile of IRF3 (30 minutes to 2hr) rather than of ISGF3 (2hr to 6hr) (Figure 3.5). Furthermore, the p50ko phenotype in response to IFNβ stimulation, which only activates ISGF3, is more modest than with LPS (Appendix Figure 3.1). Interestingly, recent protein binding microarray experiments revealed that IRF3 prefers "GG"AAAC IRE sequences, whereas the ISGF3 complex prefers the IRE sequence "TG"AAAC (Badis et al., 2009, Figure S9L). Hence, we speculate that p50 homodimer bound to G-

IREs functions primarily as a competitor for IRF3 and to a lesser extent for the ISGF3 complex.

Using mathematical modeling we have begun to explore what the functional consequences of NFκB p50:p50-IRE cross-binding may be within gene regulatory circuits. Combinatorial regulation by sequence specific transcription factors is thought to form the basis for stimulus-specific gene expression (Brivanlou and Darnell, 2002; Carey et al., 1990; Ptashne and Gann, 2002). Mathematically, such regulation has been described with Boolean logic gates (Buchler et al., 2003; Mayo et al., 2006), but such studies presume stimulus-responsive transcription factors that have no or negligible basal activity. However, in mammalian cells even highly stimulus-responsive transcription factors such as NFκB and IRF3/ISGF3 show detectable basal activities; indeed, Western blots of IRF3 indicate that as much as 30% may be constitutively nuclear (Appendix Figure 3.8), though not all of it is likely active. A thermodynamic formulation of an AND gate describes the dose response behavior of NFκB and IRF and reveals that such basal activity renders the AND gate leaky; specifically, stimuli such as CpG that activate NFκB but not IRF, will nevertheless result in appreciable AND gate promoter activity. In this context, the constitutive repressor NFκB p50:p50 binding to IREs functions as a gate keeper of combinatorial AND gate promoters. In addition to the presumed synergistic interactions of coordinated transcription factors, our combined computational-experimental study of interferon-β confirms that such a gating mechanism is critical for

enforcing the stimulus-specificity in gene expression and the subsequent antiviral responses.

More broadly, in the context of basal activities of transcription factors, the mechanisms that restrict transcriptional responses to appropriate stimuli or enforce strict synergy requirements remain incompletely understood. Chromatin regulation provides a means for regulating the accessibility of transcriptional activator binding sites and has also been proposed as providing thresholding functions (Lam et al., 2008). However, the thresholding function of competitive binding by the p50:p50 homodimer may not introduce a delay that chromatin opening steps may require. In addition, it may be tunable, as the homeostatic abundance of p50 may be regulated to control the responsiveness of the IFN program. Indeed, prolonged NFκB activation results in elevated p50 levels and elevated p50:p50 binding to IREs (Figure 3.5) and this may restrict the IFN response further to a narrow set of stimuli, without affecting the kinetics of activation.

Understanding the role of homeostatic thresholding factors in stimulus-responsive gene control is critical in developing therapeutic strategies based on manipulating gene expression responses. In harnessing the interferon response for anti-viral and anti-cancer treatment, our study suggests that the cross-regulatory functions of the NFκB p50 homodimer may limit efficacy of the interferon treatment. However, an understanding NFκB p50 homodimer's promoter gating function holds promise for tuning the specificity of adaptive immune adjuvants and innate immune priming strategies.

MATERIAL AND METHODS

Cell culture. BMDMs were isolated from C57BL/6, *nfkb1*$^{-/-}$, *ifnar*$^{-/-}$ and *p105*$^{-/-}$ mice and cultured in L929 cell-conditioned medium for 5-8 days. Primary MEFs were prepared with E12-E14 embryos from C57BL/6 and *nfkb1*$^{-/-}$ mice and cultured in DMEM containing 10% BCS for 5-6 passsages. TLR4-HEK293T cells were stably transfected with the human TLR4a, MD2 and CD14 genes (InvivoGen). Cells were stimulated with 0.1ug/ml LPS (Sigma, B5:055), 100 nM CpG (ODN 1668, InvivoGen) and 10 (BMDMs) or 2500 (MEFs) U/ml murine IFNβ (Biogen, Inc). A higher concentration of IFNβ was used in MEFs because MEFs exhibit lower IFN responsiveness than macrophages.

Transcriptome and Bioinformatic analysis. RNA were extracted with Qiagen RNeasy kit and hybridized to Illumina mouse RefSeq Sentrix-8 V1.1 and V2 BeadChips or Codelink mouse Uniset 1 (GE Healthcare) microarrays at the UCSD Biogem facility. De novo motif search were performed with the promoter sequences 1kb upstream and 0.5kb downstream of the transcription start site using the motif search program Homer, developed by Dr. Chris Benner (Heinz et al.). An in-depth description and benchmarking of this software suite can be found at http://biowhat.ucsd.edu/homer/.

RNA from littermate wt and *nfkb1*$^{-/-}$ MEFs stimulated with 0.1 μg/ml LPS for 0, 1, 3 and 8 hour, *crel*$^{-/-}$*p65*$^{-/-}$ MEFs stimulated with 0.1 μg/ml LPS for 0, 1, 3 and 8 hour and wt MEFs stimulated with 0.1μg/ml LPS (3 biological repeated samples), 10 ng/ml

TNF, 50 ng/ml PDGFBB, 2500 U/ml IFNβ for 0, 1, 3 and 8 hour were hybridized to Illumina mouse RefSeq Sentrix-8 V1.1 BeadChips arrays. Probes with ≥ 2 expression fold change in any stimulus conditions in wild-type cells were selected. K-means clustering was performed with selected probes on all of the wt samples (TNF, IFNβ, PDGFBB and 3 LPS datasets) and *crel*$^{-/-}$*p65*$^{-/-}$ samples stimulated with LPS. Details of this analysis will be published elsewhere (C.S.C. and A.H. in preparation). The *nfkb1*$^{-/-}$ LPS dataset was matched to the final clusters and the average expression fold change for each gene cluster was calculated and shown next to the heatmap in Figure 3.3.

RNA from littermate wt and *nfkb1*$^{-/-}$ BMDMs stimulated with 0.1μg/ml LPS or 10 U/ml IFNβ or 100 nM CpG for 1, 3 and 8 hour were hybridized to Illumina mouse RefSeq Sentrix-8 V2 BeadChips arrays. Raw expression data were normalized using several unstimulated control data sets. Probes with ≥ 2^1.2 expression fold change at any timepoint (1, 3, 8 hrs) of the LPS timecourse were selected. Expression fold change values from multiple probes for a single gene (accession number) were averaged. K-means clustering was performed with wt (LPS), p50ko (LPS), wt (IFNβ), *infar*$^{-/-}$ (LPS) time-course datasets and shown in Figure 3.4, Appendix Figure 3.1, and Figure 3.25.

BMDM p50ko expression phenotype and "G-rich" IRE classification: Microarray data (Codelink mouse Uniset 1 (GE Healthcare) microarrays) derived from wt and *nfkb1*$^{-/-}$ BMDMs stimulated with 0.1 μg/ml LPS for 6 hour (two biological repeats), wt BMDMs stimulated with 100 U/ml IFNβ and *infar*$^{-/-}$ BMDMs stimulated with 0.1 μg/ml LPS. Genes with ≥ 2 fold expression change in any condition were

subjected to K-means clustering. 107 genes that were IFNβ inducible and showed IFNAR-dependence when induced by LPS were selected and shown in Supplementary Table 1. Their p50ko expression phenotype in response to LPS (Figure 3.10B) was determined using replicate microarray data from wt and *nfkb1*$^{-/-}$ BMDMs stimulated with 0.1 μg/ml LPS and statistical analysis with Vampire(Hsiao et al., 2005). IREs were classified into G-rich or non-G-rich by examining the IRE binding sites 1kb upstream and 0.5kb downstream of the transcription start site. Sequences were downloaded from the UCSC Genome Browser (Mouse Feb. 2006 assembly; NCBI36/mm8) using the UCSC known genes track, which are based on the annotation from UniProt, RefSeq and GenBank mRNA. IRE sites were identified using the software MatInspector (Release professional 7.7.3, February 2008) (Cartharius et al., 2005) with the following parameters: Matrix Family Library Version 7.0 (October 2007), the family of matrices V$IRFF (includes the matrices V$ISRE.01, V$IRF1.01, V$IRF2.01, V$IRF3.01, V$IRF4.01, V$IRF7.01), and the optimized matrix similarity threshold (the number of false positive matches found in non-regulatory test sequences is minimized) of 0.88. IREs containing GGGN or GAGG or GTGG motifs within 7 bp from the 5' end of the IRE element ("G/A/T"AAA or "A"ANNGAAA) were considered to be "G-rich" IREs, as indicated in Supplementary Table 1.

Biochemical Assays.

Immunoblots and EMSAs were conducted using standard methods as described (Basak et al., 2007; Werner et al., 2005). Total RNA was isolated using Qiagen RNAeasy kit from BMDMs or MEFs treated as indicated. RNA was reverse transcribed with Superscript II reverse transcriptase (Invitrogen) and resulted cDNA was used for real-time qPCR analysis (SYBRgreen). ChIP assay was conducted as previously described (Ogawa et al., 2004; Perissi et al., 2004). Briefly, BMDMs were stimulated and fixed with 1% formaldehyde. Cross-linked pellets were sonicated to obtain fragments of 300-800bp and incubated with antibody overnight. Immunoprecipitated DNA fragments were reverse cross-linked, purified and amplified by qPCR. Reporter assays were conducted by transfecting TLR4-HEK293T cells in 24-well plates with 40ng of indicated IRE luciferase reporter construct and 40, 120 or 300 ng of p50 or p50 DBD mutant (R56A,Y57A) plasmid construct using Lipofectamin 2000 (Invitrogen). 48 hrs after transfection, cells were stimulated with LPS (0.1 μg/ml) for 6hrs and luciferase units were measured by standard methodology using the Dual-Luciferase Reporter Assay System (Promega). IRE luciferase activities were normalized to β-galactosidase activities measured by the Galacto-Light Plus System (Applied Biosystems).

Antibodies used in immunoblots, super-shift assays and ChIP assays were anti-p65 (Santa Cruz Biotechnology, sc-372), anti-p50 (sc-114), anti-Stat2 (sc-950), anti-a-tubulin (sc-5286), anti-p50 (Dr. Nancy Rice, NC-1263), anti-Lamin A/C (Cell Signaling, #2032) and anti-IRF3 (Cell Signaling, #4962).

EMSA probes:

κB (Werner et al., 2005):

GCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGG

IRE(Braganca et al., 1997; Goriely et al., 2006; Park et al., 1999):

GATCCTCGGGAAAGGGAAACCTAAACTGAAGCC

Gbp1: ACAGGTGGGTGGGGGAAAAAGAAAATGAAAGGAAA

Stat2: AACCGTGAGTTCTAGGGAAAGGAAACTGAAACCAG

Ifit3: GTTTGTGAAAGGGAAGAGGAAAGTAGAAACTGAAA

ISG15: CCGTAGGAAAAGGAAACCGAAACAGAAAATAGCTC

ADAR1: AGCCTTTTCAAGGAAACGAAAGTGAACTCTGGGGA

Ccl12: GAAGAACATCTTTATGGAAGAAAGGAAACTAGAAG

IFNb (Panne et al., 2007):

TAAATGACATAGGAAAACTGAAAGGGAGAAGTGAAAGTGGGAAATTCCTCT

GAATAG

IFNb_ΔC:

TAAATGACATAGGAAAACTGAAAGGGAGAAGTGAAAGTTCTAAATAGATCT

GAATAG

IFNb_ΔAB:

TAAATGACATATCTAAACTGAAATCTAGAAGTGAAAGTGGGAAATTCCTCTG

AATAG

IFNb_ΔABC:

TAAATGACATATCTAAACTGAAATCTAGAAGTGAAAGTTCTAAATAGATCTG

AATAG

IFNb_ΔAC:

TAAATGACATATCTAAACTGAAAGGGAGAAGTGAAAGTTCTAAATAGATCTG

AATAG

IFNb_ΔBC:

TAAATGACATAGGAAAACTGAAATCTAGAAGTGAAAGTTCTAAATAGATCTG

AATAG

qPCR primers for ChIP assays:

IkBa.f: GCTTCTCAGTGGAGGACGAG

IkBa.r: CTGGCAGGGGATTTCTCAG

A20.f: CCCGGAGAAACTCCTAGGTC

A20.r: CACATGGATGTGACGTGGAA

Gbp3.f: CAAAGCTGGTTCATGTCAGG

Gbp3.r: AAGCCCTTTCTCCTCCCTTT

Tap1.f: GTCTCAGAAGGAGGCGTGTC

Tap1.r: GAGCTGGTGGAGCTGACTAGA

Il15.f: CAGAGACTGTACCGGGAGGA

Il15.r: CAGCAGGTCAAGGGTTGTTT

Stat2.f: AGGTCCCACCCTTTCTATGG

Stat2.r: CTGATTTACCCGAACCGAAC

Gbp2.f: AGCTAGCTGATTTCCCAGCA

Gbp2.r: GGAAGGAGGGAGGAAGAAAA

Tyki.f: ACTGGAGGTCTCAGCCACAG

Tyki.r: GGTGACCACAGACCCAGGAT

Trim21.f: GCCACTCAACTCCTCACTCC

Trim21.r: AGCCTGCAGTCCAAATTTCA

Intergenic.f (chr12:57147085-57147178): GATATTCCCTGTCCCCCAGT

Intergenic.r: AAGGAGCAGAGAGAGCAGTG

Ifi47.f: ATGCTGCAGGGGAAACAAAG

Ifi47.r: AAGCAATGAGCCCTAGCAGA

Dusp28.f: AGATTGGCTAGTGGGGAATG

Dusp28.r: CACTTTCATGGCTATGATTTGC

Tcea1.f: CCTCACAAGGAAATTGAAGG

Tcea1.r: CTCCCCAGGGTAACAGTGAA

Cd274.f: TCGACAGCCTCTCAGTAGCA

Cd274.r: CAGTGGCAGGTGAGTCTCTG

Lgals3bp.f: GGTGGTGGTAGTTTTGTTGTTG

Lgals3bp.r: AGCCTGCCTCAAAGGAAAAA

qPCR primers for RT-qPCR assays:

Gbp1.f: CGGAAAGAGTTAATGGCAGAGC

Gbp1.r: GTTGCAAGCTCTCATTCTGG

Gbp3.f: GATGGAGAGAGAGCCATAGCA

Gbp3.r: CCTTCTGTCTCTGCCTCAGC

Mx1.f: GACCCTGAAGGGGATAGGAC

Mx1.r: CTTGCCTTCAGCACCTCTGT

Gbp1.f: CGGAAAGAGTTAATGGCAGAGC

Gbp1.r: GTTGCAAGCTCTCATTCTGG

Ifit3.f: CCAGCAGCACAGAAACAGAT

Ifit3.r: GAAATGGCACTTCAGCTGTG

Ifit2.f: CGCTTTGACACAGCAGACAG

Ifit2.r: GTCGCAGATTGCTCTCCAGT

Vig1.f: GCTGGAAGGTTTTCCAGTGC

Vig1.r: CTTCCCTCAGGGCATCTTCT

Tyki.f: AGACAGGTACTGGCATAGCACA

Tyki.r: ACTGTAGGCCTCCACTCACC

Gapdh.f: AACTTTGGCATTGTGGAAGG

Gapdh.r: GGATGCAGGGATGATGTTCT

"G-rich" and "non-G-rich" IRE sequences in luciferase reporter constructs:

G-rich IRE: TCGACGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAGGGAAAG

Non-G-rich IRE: TCGACAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAG

Computational Modeling.

Thermodynamic formulations of promoter activity *f* are shown in the figures and based on prior work (Bintu et al., 2005a; Bintu et al., 2005b). Thermodynamic models of promoter-transcription factor interactions employed parameter values that were based on measurements and reasonable estimates. The nuclear p50 homodimer concentration was measured to be about 30,000 dimers (Fig. 1B). Based on macrophage cell and nucleus sizes measured by us and others (Garrick et al., 1986; Swanson et al., 1991) to be 1 - 2 pl and 0.15 - 0.3 pl respectively, this translates to a molar concentration of at least 200 nM. To ensure that our modeling results are conservative, we have also considered lower p50 homodimer concentrations. The stimulus-induced nuclear NFκB concentration has previously been determined to be peak at about 100 nM (Werner et al., 2005). Based on that the IRF concentration is assumed not to exceed 100 nM, with a basal level 5-10% of that. The dissociation constants were reasonably estimated based on simple guidelines: NFκB p50:p65 has higher affinity than p50:p50 for κB sites (i.e. $K_{NF\kappa B\text{-}\kappa B} < K_{p50\text{-}\kappa B}$), IRF has higher affinity than p50:p50 for IREs (i.e. $K_{IRF\text{-}IRE} < K_{p50\text{-}IRE}$), NFκB p50:p65's affinity for κB sites is higher than IRF's affinity for IREs (i.e. $K_{NF\kappa B\text{-}\kappa B} < K_{IRF\text{-}IRE}$), and p50:p50 has higher affinity for κB sites than IREs (i.e. $K_{p50\text{-}\kappa B} < K_{p50\text{-}IRE}$).

The specific parameter values were as follows: For Figure 4E and 4F, [p50:p50] = 100 nM, [IRF] = 1-100 nM, $K_{IRF\text{-}IRE}$= 10 nM, $K_{p50\text{-}IRE}$= 100 nM (Figure 4E) or 10-1000 nM (Figure 4F). For Figure 5A and 5B, [p50:p50] = 100 nM, [IRF] = 0-50 nM,

[NFκB/p65:p50] = 0-50 nM, $K_{IRF\text{-}IRE}$= 5 nM, $K_{p50\text{-}IRE}$= 10 nM, $K_{p50\text{-}\kappa B}$= 5 nM, $K_{NF\kappa B\text{-}\kappa B}$= 1 nM. For Figure 5C, parameter sensitivity analyses were performed by varying the binding dissociation constants $K_{IRF\text{-}IRE}$, $K_{p50\text{-}IRE}$, $K_{p50\text{-}\kappa B}$, and $K_{NF\kappa B\text{-}\kappa B}$ from 1 to 1000 nM as indicated in the figure. Hill coefficient for the term [IRF/ $K_{IRF\text{-}IRE}$] was also varied from 1 to 3.

The kinetic model of IFNβ mRNA (*r*) expression was modeled by an ODE,

$$\frac{dr}{dt} = k_i - k_d r \quad ,$$

where k_i is the mRNA synthesis rate constant (in arbitrary mRNA concentration units min^{-1}) and k_d is the mRNA degradation rate constant (in min^{-1}). For Figure 6B and 6C, the ordinary differential equations model has the following parameters $K_{IRF\text{-}IRE}$ = 5 nM, $K_{p50\text{-}IRE}$= 10 nM, $K_{p50\text{-}\kappa B}$= 5 nM, $K_{NF\kappa B\text{-}\kappa B}$ = 1 nM, [p50:p50] = 100 nM. The time-dependent inputs of the model are: basal [NFκB] = 1 nM, and in response to LPS and CpG it is 10, 50, 40, 30 nM at 30, 120, 360, 720 minutes, respectively; basal [IRF] = 7.5 nM and in response to LPS it is 20, 50, 40, 20 nM at 30, 120, 360, 720 minutes, respectively and were estimated from gel shift experiments and nuclear western blots. The mRNA halflife was estimated at 1 hour resulting in a degradation rate constant k_d = 0.0116 min^{-1}. The mRNA synthesis rate constant k_i was arbitrarily set to 0.24 AU min^{-1}. Simulations were done with the numerical solver of the Matlab ode23 function.

Viral infection.

Prior to infection, BMDMs were seeded into 24-well plates, allowed to adhere and treated with indicated stimuli for 24 hours. BMDMs were infected with live MCMV-GFP and were analyzed for GFP expression after 48 hours by flow cytometric analysis.

ACKNOWLEDGEMENTS

I thank James Lee for the initial finding on p50 binding to IRE probes, Kristyn Feldman for the anti-viral assays, Dr. Shilpi Verma and Dr. Chris A. Benedict for providing viral assay training for Kristyn Feldman, Dr. Gourisankar Ghosh and Dr. De-Bin Huang for the structural model, Kim Huynh for the mutation EMSAs, Dr. Mikyoung Chang and Dr. Shao-Cong Sun for the $p105^{-/-}$ mice and Dr. Julia Ponomarenko for suggestions on better documentation of genome databases. I thank Dr. Gary Hardiman and the Biogem core facility and Dr. Chris Glass for gene expression profiling; Dr. Chris Benner for his motif search program (Homer); Dr. Edward Chiang for TLR4-HEK293T cells; Dr. Masataka Asagiri and Dr. Tom Huxford (SDSU) for critical review of the manuscript; Drs. Terry Hwa, Victor Nizet and Chang Chang (U Penn), as well as members of the laboratory, Vivien Wang and Dr. Ville Veckman for helpful discussions and support.

Chapter 3, in full, is a reprint of the material as it appears in Science Signaling 2011**.** Cheng CS, Feldman KE, Lee J, Verma S, Huang DB, Huynh K, Chang M, Ponomarenko JC, Sun SC, Benedict CA, Ghosh G, Hoffmann A,

American Association for the Advancement of Science (AAAS), 2011. The thesis author was the primary investigator and author of this paper.

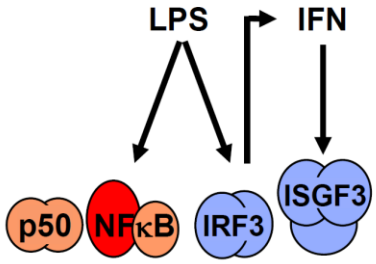


Figure 3.1 LPS- and IFNβ-induced transcription factors

A schematic of LPS- and IFNβ-induced activation of transcriptional activators (NFκB, IRF3 and ISGF3) and the constitutive transcriptional repressor (p50:p50).

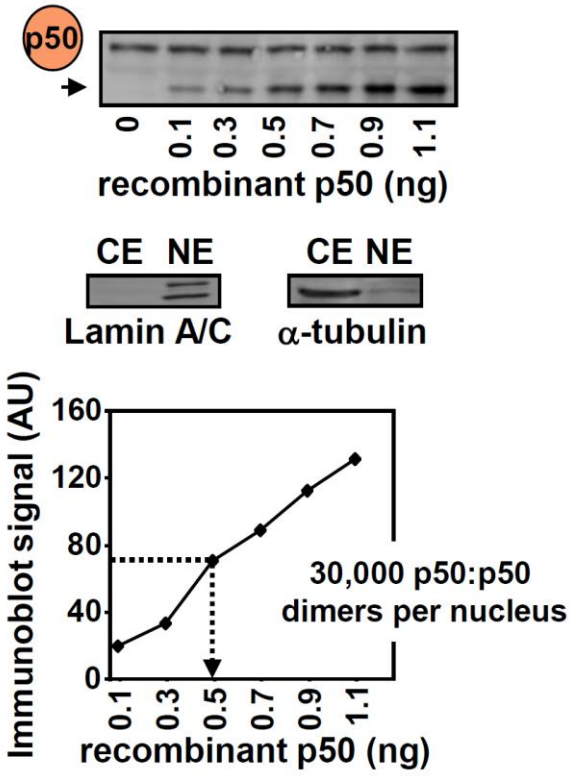


Figure 3.2 Quantitation of the basal nuclear p50 protein amount

Immunoblot of a nuclear extract prepared from non-treated wt BMDMs mixed with increasing amounts of recombinant p50 (amino acid 1-325) protein (0.1 to 1.1 ng). Nuclear and cytoplasmic fractionation was confirmed by Lamin A/C and α-tubulin immunoblots (lower panel). Signals of p50 immunoblot were quantitated and graphed on a standard curve, leading to an estimate of 30,000 p50:p50 homodimers per nucleus.

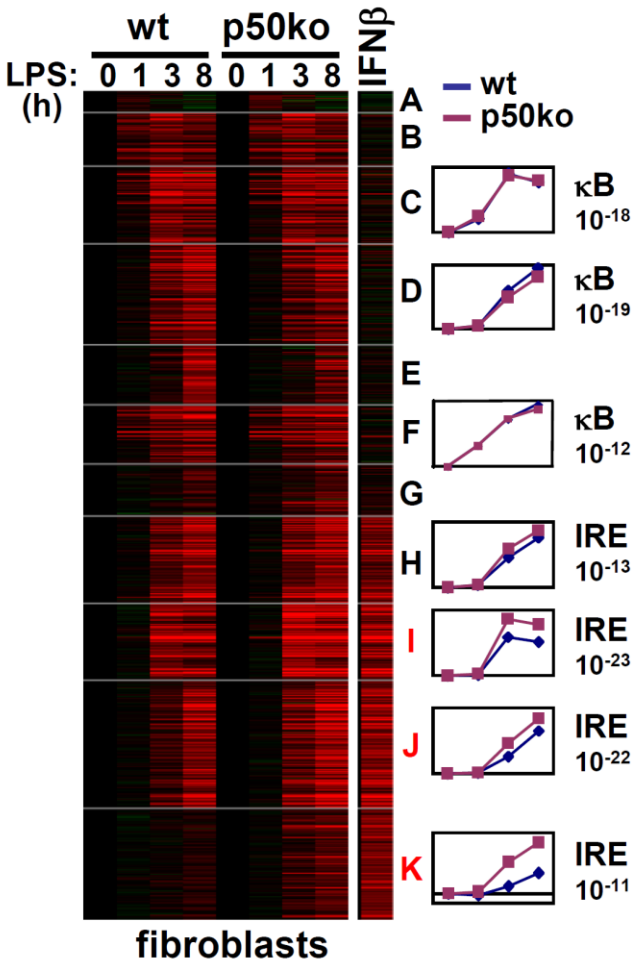


Figure 3.3 Microarray mRNA expression data from wild-type ("wt") and nfkb1-/- ("p50ko") littermate-derived MEFs

MEFs were stimulated with LPS (0.1 µg/ml) and IFNβ (2500 U/ml) was analyzed by K-means clustering as described in Supplementary Methods. Red represents stimulus-responsive gene induction, green represents repression. Cluster identifiers are indicated at the right, with red symbols indicating clusters with elevated expression in the p50ko cells. Average fold induction (log_2) of each cluster in wt (blue) and p50ko (purple) cells were graphed at 0, 1, 3 and 8 hour. Most highly enriched motifs identified *de novo* within -1.0 kb to +0.3 kb of transcriptional start sites are shown for in each cluster with p-values to indicate statistical significance.

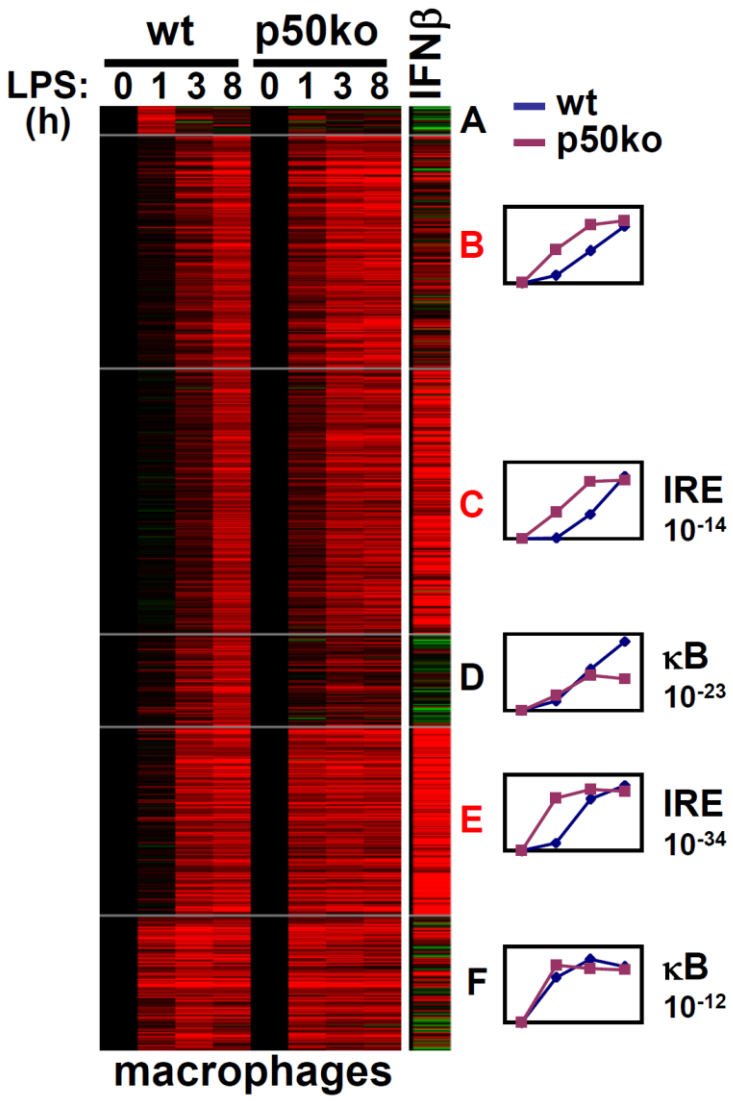


Figure 3.4 Microarray mRNA expression data from wild-type ("wt") and nfkb1-/- ("p50ko") littermate-derived BMDM

BMDMs were stimulated with LPS (0.1 μg/ml) and IFNβ (10 U/ml) was analyzed by K-means clustering as described in Figure 3.3.

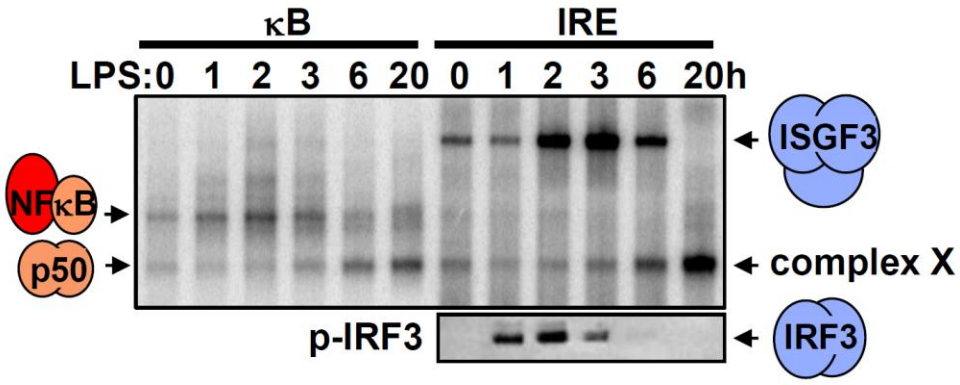


Figure 3.5 Transcription factor activation profiles in macrophages

Macrophages (RAW264.7 cells) were stimulated with LPS (0.1 µg/ml) for the indicated times. κB-binding activities (p50:p65 and p50:p50; left panel) and IRE-binding activities (ISGF3 and complex X; right panel) on an IRE probe (GATCCTCGGGAAAGGGAAACCTAAACTGAAGCC) were revealed by EMSA, and an immunoblot (middle) shows the activation profile for nuclear phospho-IRF3 (p-S396).

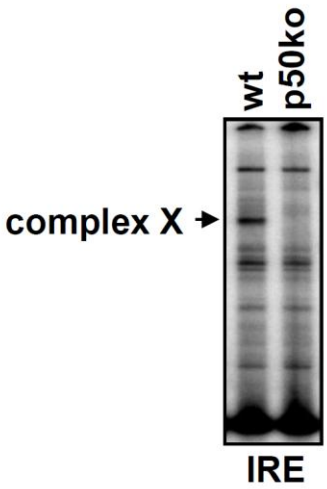


Figure 3.6 IRE-binding activities contained in nuclear extracts derived from wt and nfkb1-/- ("p50ko") BMDMs

BMDMs stimulated with 0.1µg/ml LPS for 24 hour were resolved by EMSA.

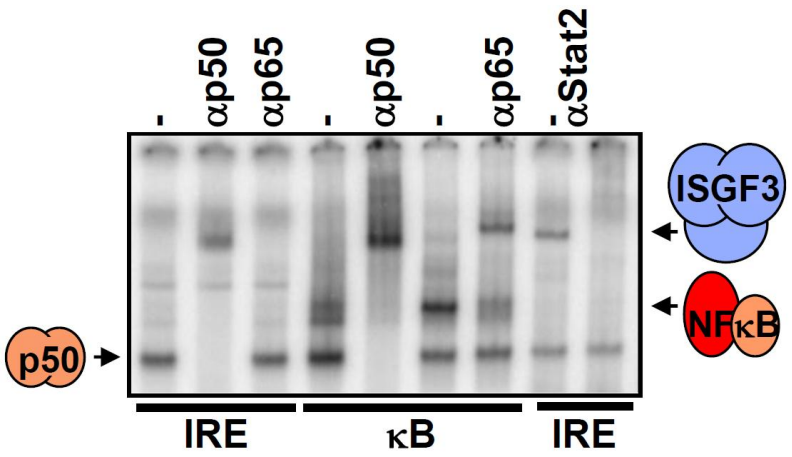


Figure 3.7 The composition of the novel complex X

Complex X was examined by supershift analysis using indicated antibodies on the IRE probe, and mobility comparison with complexes bound to an HIV κB site probe.

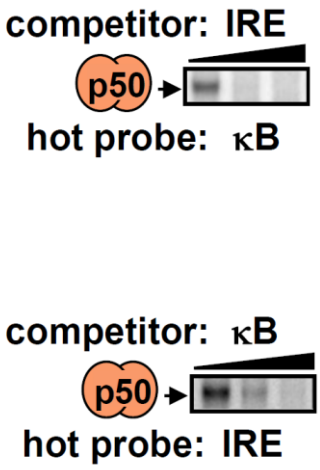


Figure 3.8 Competition assays with IRE and κB probes

DNA binding site competition assays (10x or 200x excess) with labeled κB probe and unlabelled IRE probe (top panel) or labeled IRE probe and unlabelled κB probe (bottom panel) and nuclear extracts from macrophage cells (RAW264.7) stimulated with LPS for 24 hour.

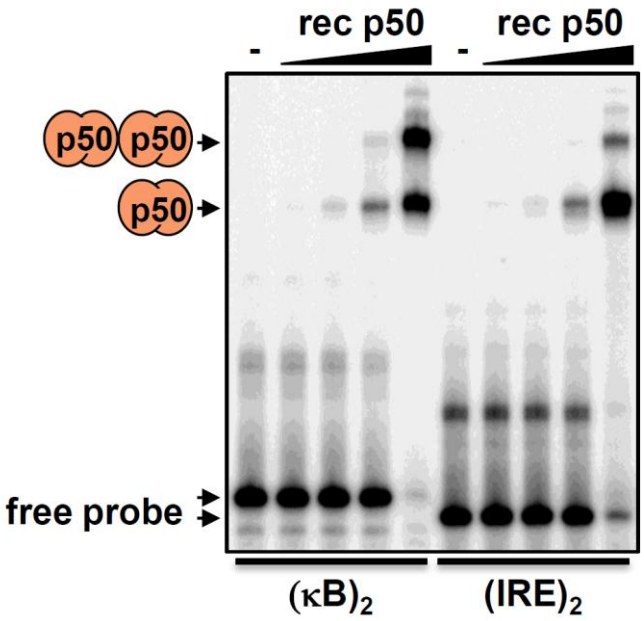


Figure 3.9 Comparison of p50:p50 DNA binding affinities to κB and IRE probes
EMSA to measure relative binding affinity of p50:p50 for κB or IRE probes. Increasing amounts of recombinant p50 (1, 10, 100, 100 ng) was reacted with a κB site probe containing two κB binding sites (left panel) or an IRE probe containing 4 IRE binding sites (right panel); both probes allow binding of two p50:p50 dimers.

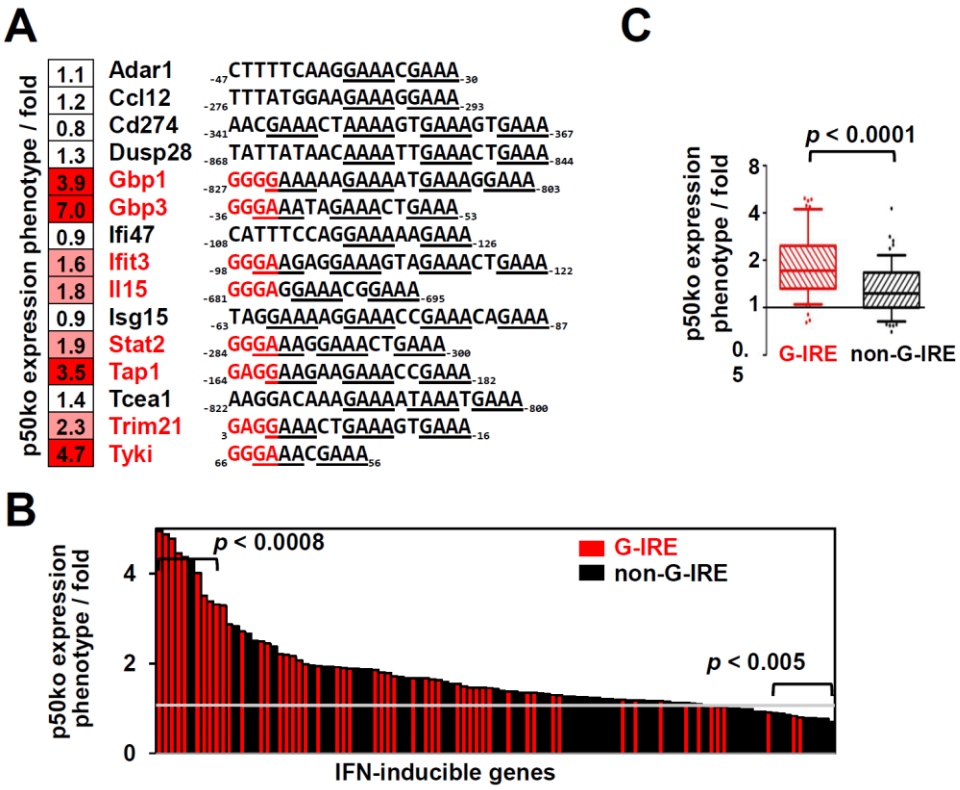


Figure 3.10 NFκB p50 binds to a subset of interferon response genes via a "G-IRE" (A) Known IRE-regulated genes distinguished as p50-repressed (red) or not (black). IRE sequences are underlined; guanine-rich sequences are highlighted in red. The p50ko expression phenotype represents the ratio of the LPS-induced (0.1 µg/ml at 6 hour) expression fold change in p50ko over wt BMDMs in duplicate microarray experiments (white; fold change < 1.5, pink: fold change ≥ 1.5 and < 3; red: fold change ≥ 3). (B) The p50ko expression phenotype (y-axis) for 107 interferon-β-inducible and IFNAR-dependent LPS-inducible genes (x-axis). Genes that contain one or more G-IREs in the -1.0kb to +0.5kb regulatory region are colored in red. Probabilities that the number of G-IRE-containing genes among the top or bottom 10% genes within distribution occurred by chance are indicated. (C) Unpaired t-test of the p50ko expression phenotype associated with G-IRE (red) and non-G-IRE (black) containing genes in (B). Edges of the boxplot represent the 25th and 75th percentiles and whiskers represent the 10th and 90th percentiles.

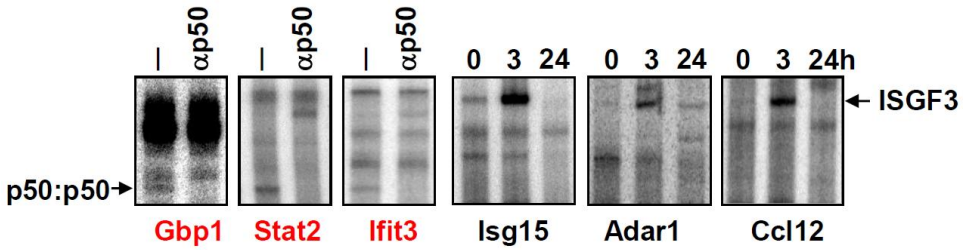


Figure 3.11 p50:p50 physically binds to the G-IRE genes

p50:p50 dimer and ISGF3 DNA binding activities were resolved by EMSA and supershift assay with p50 antibody using extracts made from macrophages (Raw264.7 cells) stimulated with LPS (0.1 µg/ml) for 24 hours using probes derived from p50-repressed genes with G-IREs (red) or probes derived from p50-independent genes containing non-G-IREs (black). LPS-stimulated macrophages contain higher levels of nuclear p50:p50 and allow for a positive control, the confirmation that these sequences can in fact bind ISGF3.

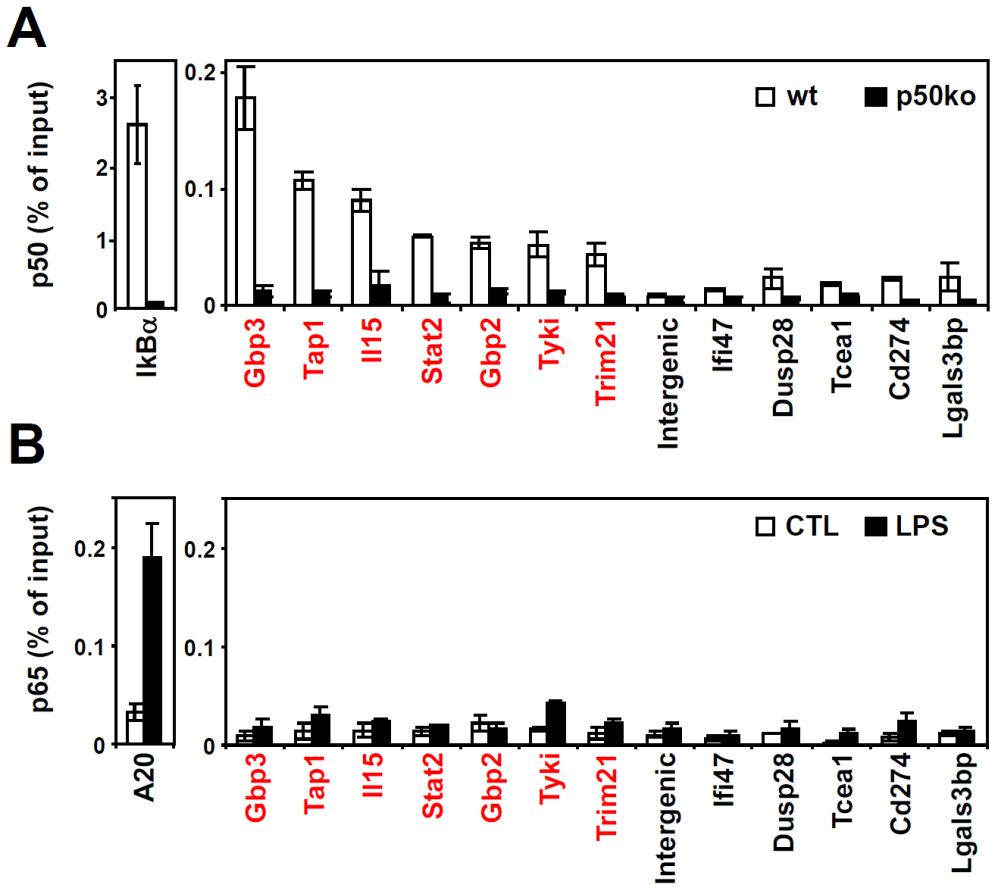


Figure 3.12 NFκB p50 binds to G-IRE promoters in vivo

(A) ChIP followed by qPCR analysis indicate recruitment of p50 to the p50-repressed (and G-rich, in red) but not to p50-independent (and non-G-rich, in black) IRE sequences in wt (white bar) and p50ko (black bar) BMDMs stimulated with LPS (0.1 μg/ml) for 24 hours. IκBα functions as a positive control for NFκB target gene. (B) ChIP followed by qPCR analysis of the recruitment of p65 on p50 repressed, G-IREs (red) and p50-independent, non-G-rich IREs (black), as well as a positive control κB target gene A20 in response to LPS (0.1 μg/ml, black bar) or PBS (CTL, white bar) for 1 hour in wt BMDMs.

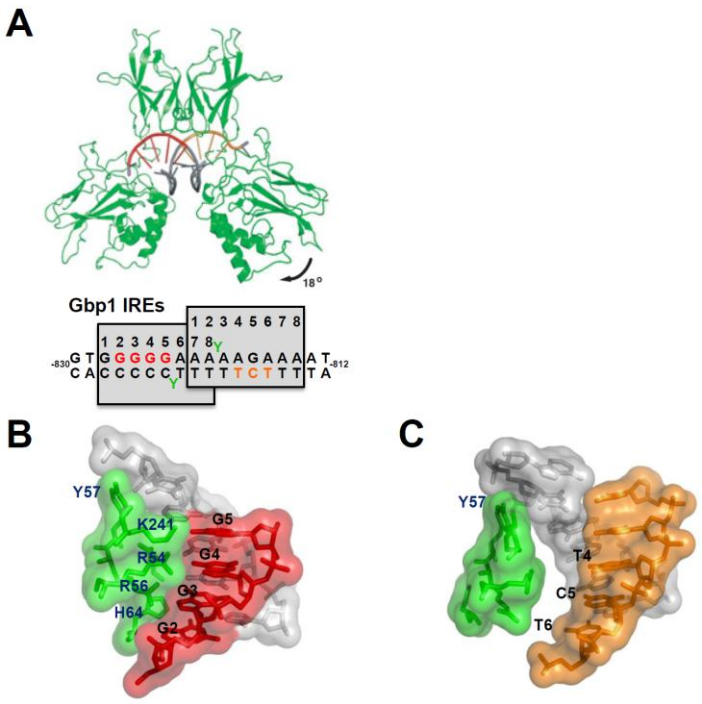


Figure 3.13 Molecular determinants of NFκB p50 homodimer binding to a G-IRE sequence

(A) Ribbon model of the structure of p50:p50 dimer on the G-IRE/IRE of Gbp1, based on prior X-ray crystallography structures on κB sites (Ghosh et al., 1995). In both the ribbon model and the G-IRE sequence below, the DNA in the vicinity of the first monomer is indicated in red, the DNA in the vicinity of the second monomer in orange. Conserved tyrosines ("Y") that contribute to affinity by inserting themselves between bases are indicated in green. Note that there is rotational flexibility between the DNA and dimerization domains, as indicated for the 2nd monomer. The DNA binding domain is shown to have moved away from the DNA 2nd half-site. (B) Space filling model detailing the amino acids within the first monomer making close contacts with the G-IRE sequence. Numbering is with respect to the center of a palindromic sequence that would apply to the κB-consensus sequence. (C) Space filling model revealing a gap between the protein surface and the second IRE sequence within Gbp1. This gap is opened by the rotational flexibility between dimerization and DNA binding domains. Note that tyrosine 57 provides for affinity as it is anchored by hydrogen bonds to the phosphate backbone and makes van der Waals contacts with the bases A2 and A3 of the second IRE.

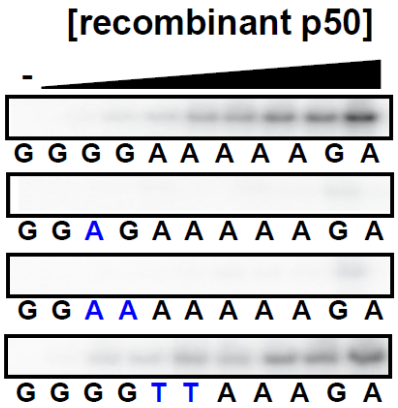


Figure 3.14 Determining the p50:p50 dimer binding specificity for IREs

EMSA was performed with a titration of recombinant p50 proteins (1, 5, 10, 20, 40, 80, 200, 500 nM) using the G-IRE-containing probe derived from Gbp1 or indicated mutant variants thereof. These experiments confirm the importance of the G-rich motif in the 5'-end of the first IRE, as opposed to central nucleotides. The EMSA was performed by Kim Huynh.

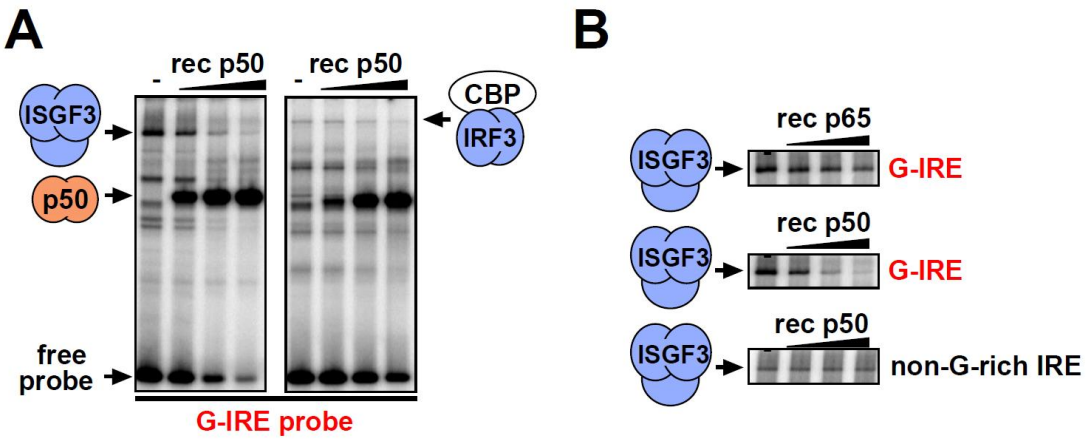


Figure 3.15 NFκB p50 homodimers repress G-IREs by competing with IRFs

(A) Competition EMSAs demonstrating that increasing amounts of recombinant p50:p50 dimers (25, 250 and 1000 nM) competes with nuclear ISGF3 complexes (contained in LPS-stimulated (0.1µg/ml, 24hrs) macrophage Raw264.7 cells) in binding to the G-IRE probe (left) or with nuclear IRF3-CBP complexes (contained in polyIC-stimulated (5.0 µg/ml, 1 hr) HeLa cells) in binding to the Gbp3 IRE probe (right). (B) Competition EMSAs demonstrating that nuclear ISGF3 complexes binding to IRE probes are not efficiently competed with recombinant p65:p65 dimers (25, 250 and 1000 nM, upper panel), but are competed by recombinant p50:p50 dimers (25, 250 and 1000 nM, middle panel), and that ISGF3 complexes binding to the non-G-rich ADAR1 IRE probe are not competed by recombinant p50:p50 dimers (25, 250 and 1000 nM, bottom panel).

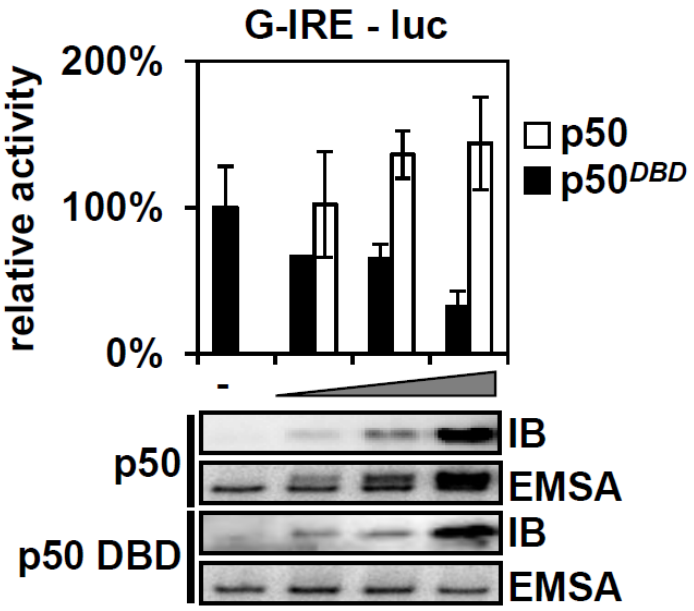


Figure 3.16 NFκB p50 represses a G-IRE reporter

p50, but not a p50 DBD mutant (R56A,Y57A) inhibits LPS-induced activation of G-IRE-containing promoters in a dose-dependent manner. TLR4-HEK293T cells were transfected with IRE reporter plasmids (40 ng) and p50 or p50 DBD mutant expression plasmid (0, 40, 120 and 300 ng) for 48 hrs in 24 well plates. Cells were stimulated with LPS (0.1 µg/ml) for 6 hours. Relative G-IRE vs. non-G-rich IRE luciferase activity was graphed as a function of promoter activity in the absence of p50. Below, immunoblot of nuclear p50 protein and p50:p50 IRE-binding activity resolved by EMSA (the lower band represents constitutively expressed endogenous p50:p50).

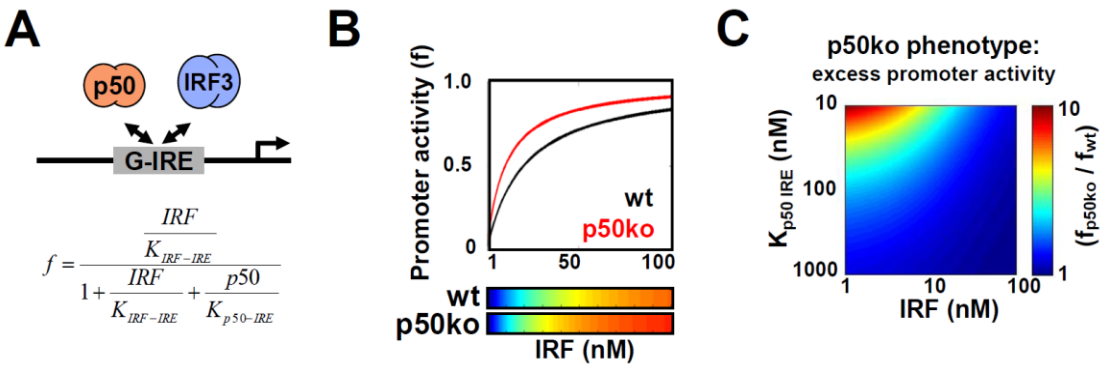


Figure 3.17 Mathematical modeling of the p50 homodimer repressing G-IRE targets genes

(A) A schematic and thermodynamic formulation of promoter activity f (Bintu et al., 2005b) of a G-IRE-containing promoter subject to competitive binding by p50:p50 dimer and IRFs. (B) The mathematical solution of promoter activity (f) in wt and p50ko as a function of the IRF concentration. The color bars represent promoter activity (f) in wt and p50ko cells ranging from 0 (blue) to 1 (red). (C) Computational determination of the p50ko phenotype defined as the fold change of G-IRE-driven promoter activity (f) between p50ko and wt cells as a function of the IRF concentration and the p50 dissociation constant for the IRE ($K_{p50\text{-}IRE}$).

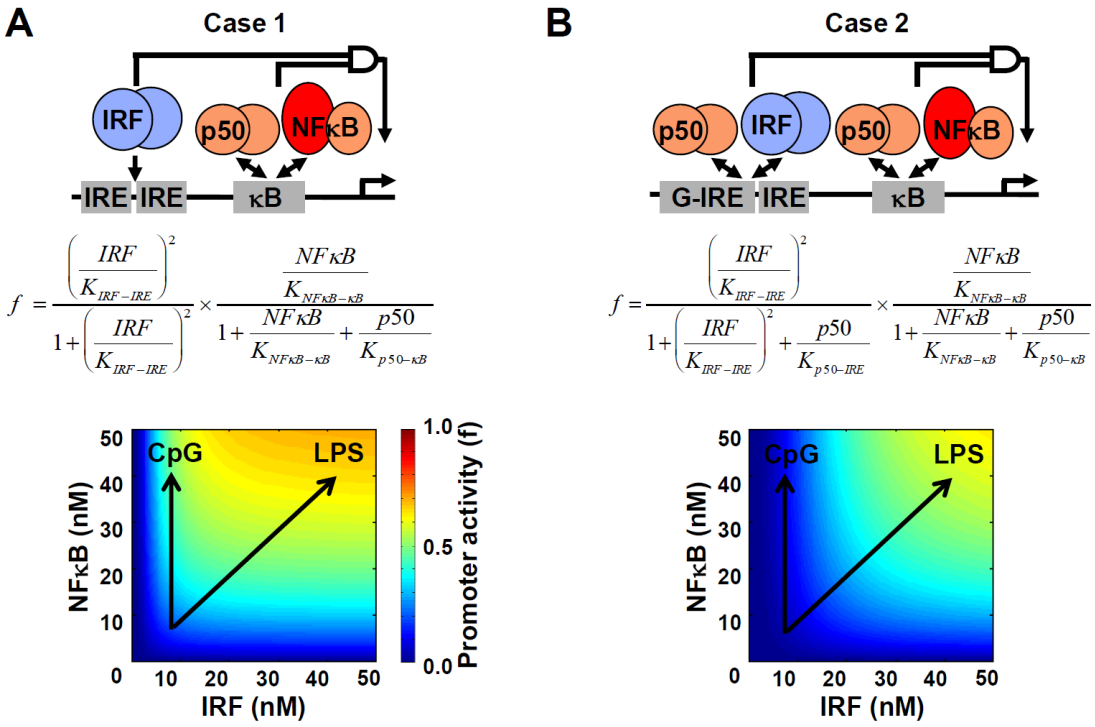


Figure 3.18 Mathematical modeling predicts that p50:p50-G-IRE interactions function to enforce stimulus-specificity of AND-gate promoters

(A) Schematic, thermodynamic formulation, and computational determination of the activity *f* of an AND-gate promoter that is a function synergistic IRF and NFκB concentrations and competitive repression of the κB site by p50:p50 homodimers. The arrows indicate how the concentrations of active NFκB alone or NFκB and IRF may increase in response to stimulation with CpG and LPS, respectively. (B) Schematic, thermodynamic formulation, and computational determination of the activity *f* of an AND-gate promoter that is a function of synergistic IRF and NFκB concentrations and competitive repression of both the G-IRE and the κB site by p50:p50 homodimers. Analogous to (A) the arrows indicate how the concentrations of active NFκB alone or NFκB and IRF may increase in response to stimulation with CpG and LPS, respectively.

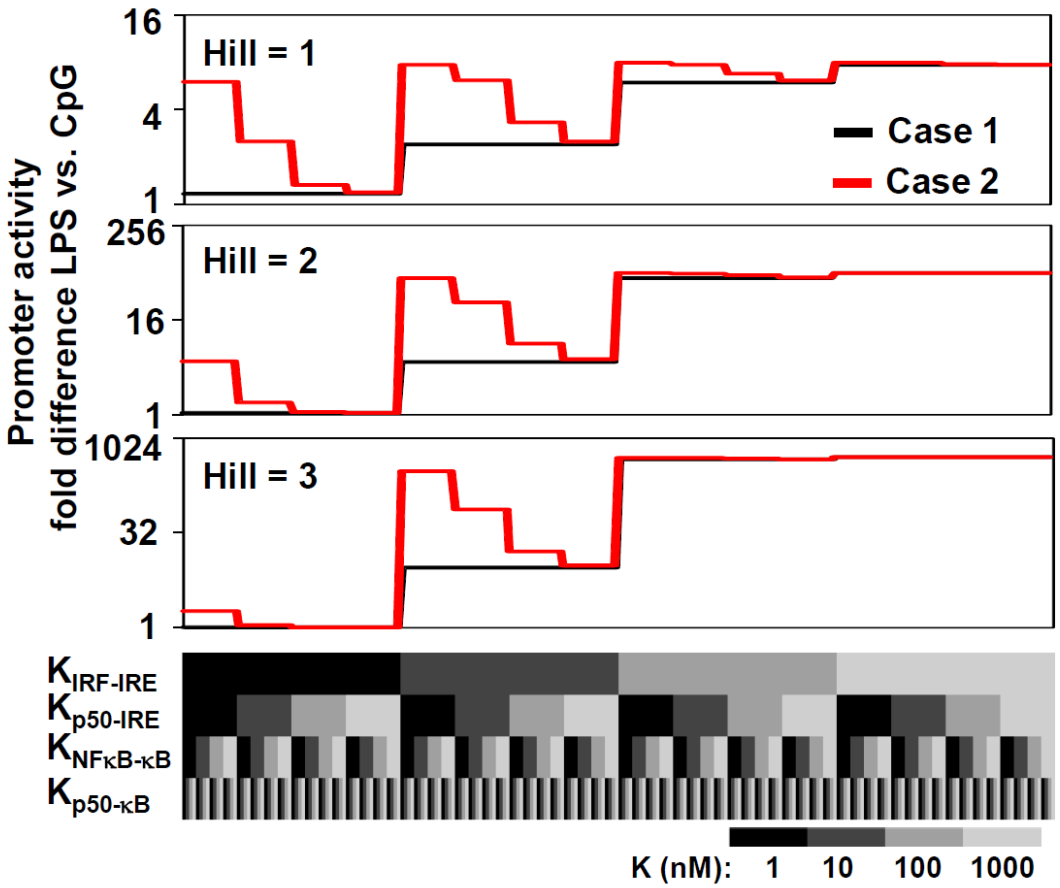


Figure 3.19 Parameter sensitivity analyses examining the relative role of the four dissociation constants (K) in determining stimulus-specificity

Dissociation constants $K_{IRF\text{-}IRE}$, $K_{p50\text{-}IRE}$, $K_{NF\kappa B\text{-}\kappa B}$ and $K_{p50\text{-}\kappa B}$ were varied from 1 to 1000 nM as indicated (bottom). The Hill coefficient for the term [IRF/$K_{IRF\text{-}IRE}$] was also varied from 1 to 3. The fold difference of LPS- vs. CpG-induced promoter activity was plotted on the Y-axis. CpG-induction involved increasing the NFκB concentration from 5nM to 40nM and LPS-induction involved increasing both IRF and NFκB from 5nM to 40nM. The data revealed that with a wide range of binding affinities and Hill coefficient values our model predicts that when p50 homodimer is bound to the IRE site (case 2, red), the stimulus-specific difference in promoter activity (LPS vs. CpG) is more pronounced than when p50 homodimer does not bind to the IRE site (cases 1, black).

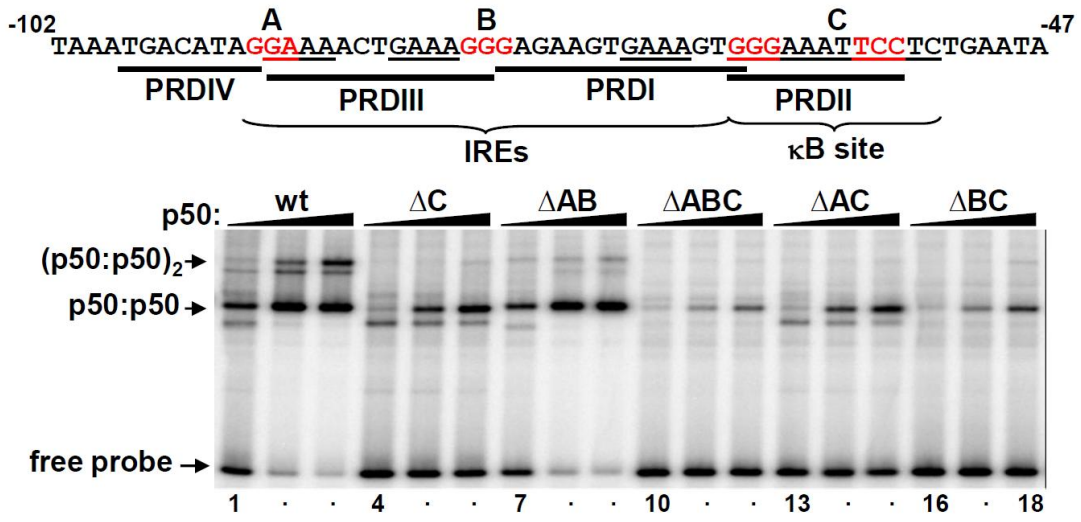


Figure 3.20 p50 homodimers bind to the IFNβ enhancer

p50 homodimer DNA-binding activity on the IFNβ enhancer (-102 to -47). Indicated IRE and κB site mutants were employed in EMSA with 10, 50 and 100 ng of recombinant p50 protein. Mutants (designated with Δ) have the indicated tri-nucleotides in red mutated to "TCT" (for GGA or GGG) or "AGA" (for TCC).

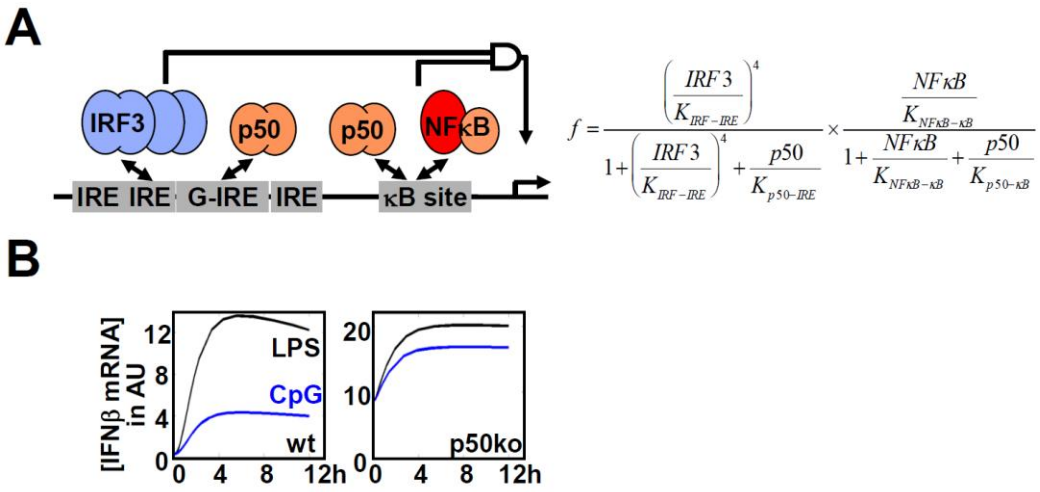


Figure 3.21 Modeling predicts stimulus specific expression of IFNβ is enforced by p50 binding to a G-IRE within its enhancer

(A) A schematic illustrating the regulation of the IFNβ enhancer by p50 homodimers on both the G-IRE and the κB site. A thermodynamic formulation of the fractional promoter activity f. (B) Computational simulations of IFNβ mRNA expression using a kinetic model driven by the fractional promoter activity f in wt and p50ko cells in response to LPS (black) or CpG (blue).

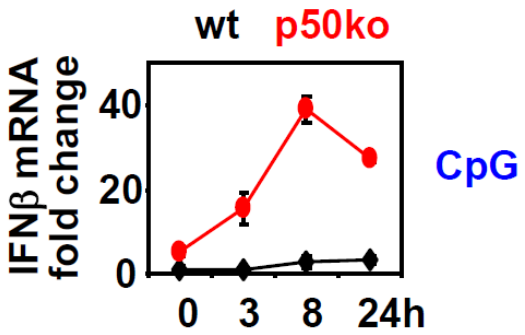


Figure 3.22 IFNβ mRNA is induced by CpG in p50ko cells

RT-qPCR determination of IFNβ mRNA expression fold change over wild-type unstimulated cells following CpG (100 nM) stimulation for indicated times of wt (black) and p50ko (red) BMDMs.

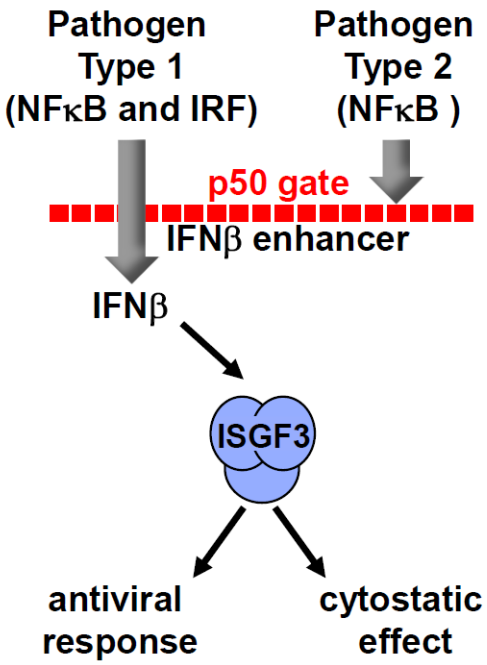


Figure 3.23 Pathogen-specific activation of the IFNβ enhancer

IFNβ enhancer gating by the p50 homodimer allows only certain pathogens (those that activate both NFκB and IRF) to activate the antiviral responses, ensuring that the cytostatic effects of IFN signaling remain shut off when anti-viral responses are not needed.

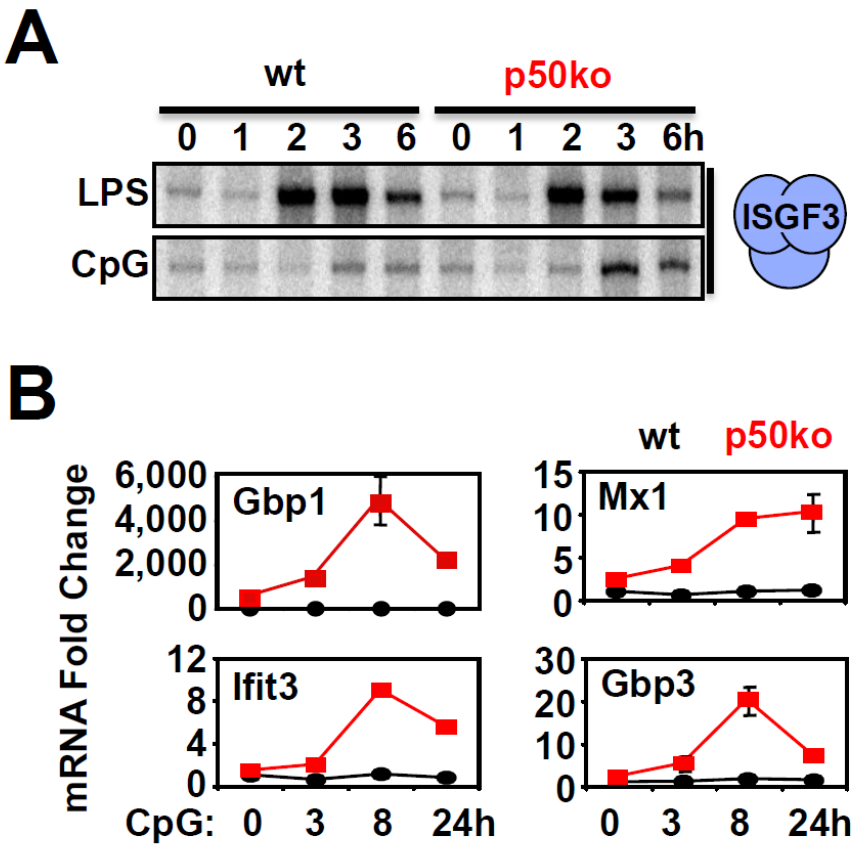


Figure 3.24 NFκB p50 restricts IFN/ISGF3 pathway and anti-viral gene expression only to IRF-inducing stimuli

(A) Nuclear ISGF3 activities were revealed by EMSA in response to LPS (0.1 μg/ml) or CpG (100 nM) using nuclear extract derived from wt and p50ko BMDMs. (B) mRNA expression fold change of IFNβ inducible genes were analyzed by RT-qPCR in wt (black) and p50ko (red) BMDMs stimulated with CpG (100 nM) at indicated times.

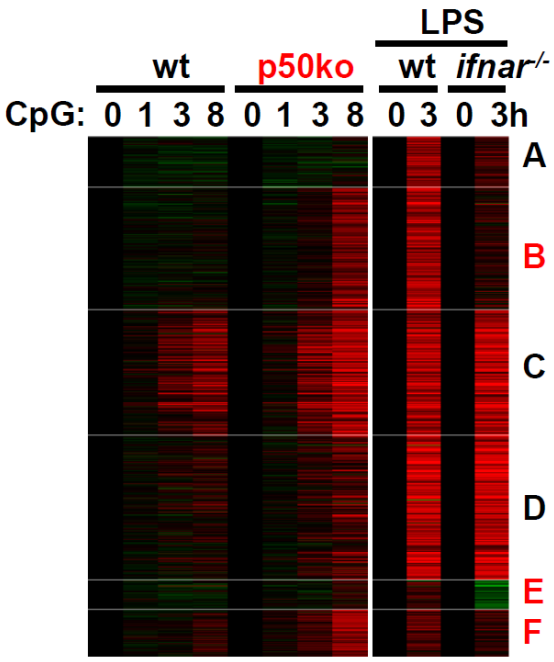


Figure 3.25 NFκB p50 inhibit anti-viral and IFNβ inducible gene expression in response to CpG

Microarray gene expression profiles of wt and p50ko BMDMs stimulated with CpG (100 nM) for 0, 3, 8 and 24 hour together with microarray expression profile of wt and *ifnar*$^{-/-}$ BMDMs stimulated with LPS (0.1 µg/ml) for 0 and 3 hour were analyzed by K-means clustering. Cluster identifiers are indicated at the right. Red indicates IFNβ/IFNAR dependent clusters.

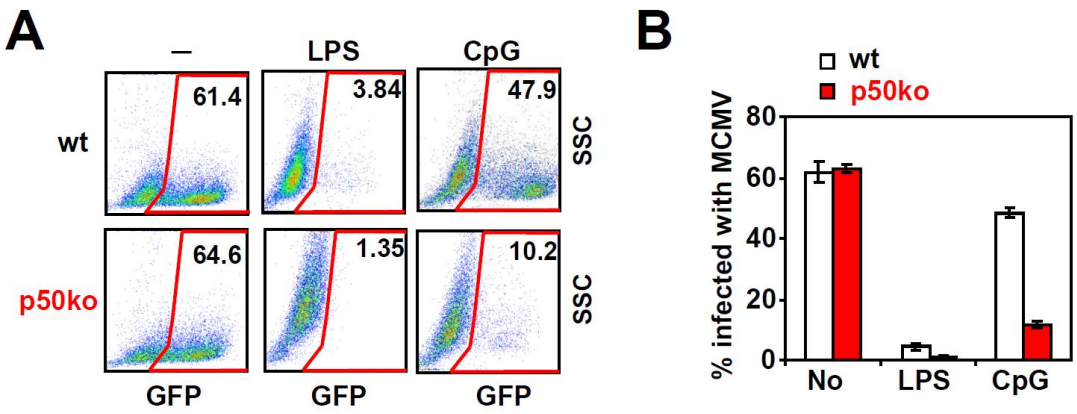


Figure 3.26 NFκB p50 restricts anti-viral response only to IRF-inducing stimuli

(A) Induction of anti-viral resistance following priming treatments. After a 24 hour treatment with PBS ("–"), LPS (0.1 µg/ml) or CpG (100 nM), wt and p50ko BMDMs were infected with MCMV-GFP. After 48 hours, productively infected cells expressing GFP were quantified (red gate) by flow cytometry (side scatter over GFP). (B) Percent of infected wt (white) or p50ko (red) BMDMs expressing GFP in (A) were graphed as average from triplicate determinations. The data are representative of three independent experiments. Kristyn Feldman performed the anti-viral assays.

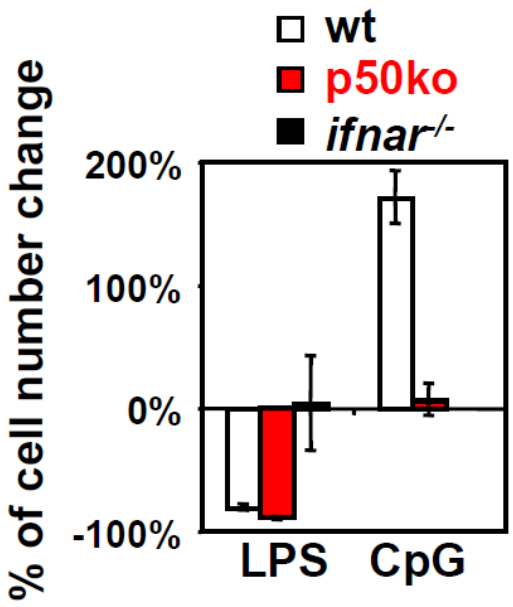
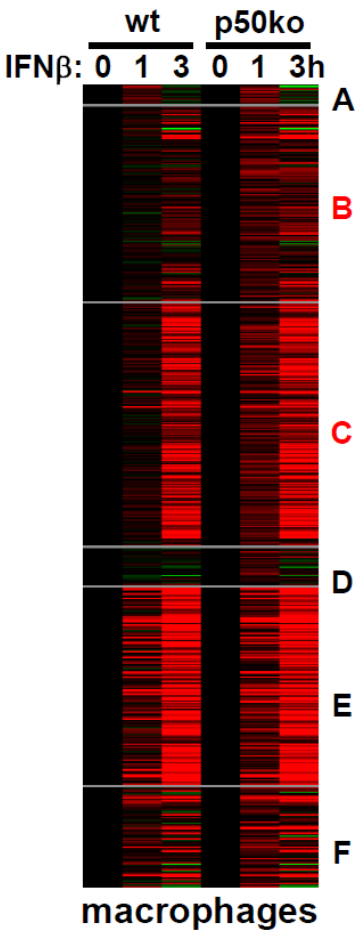


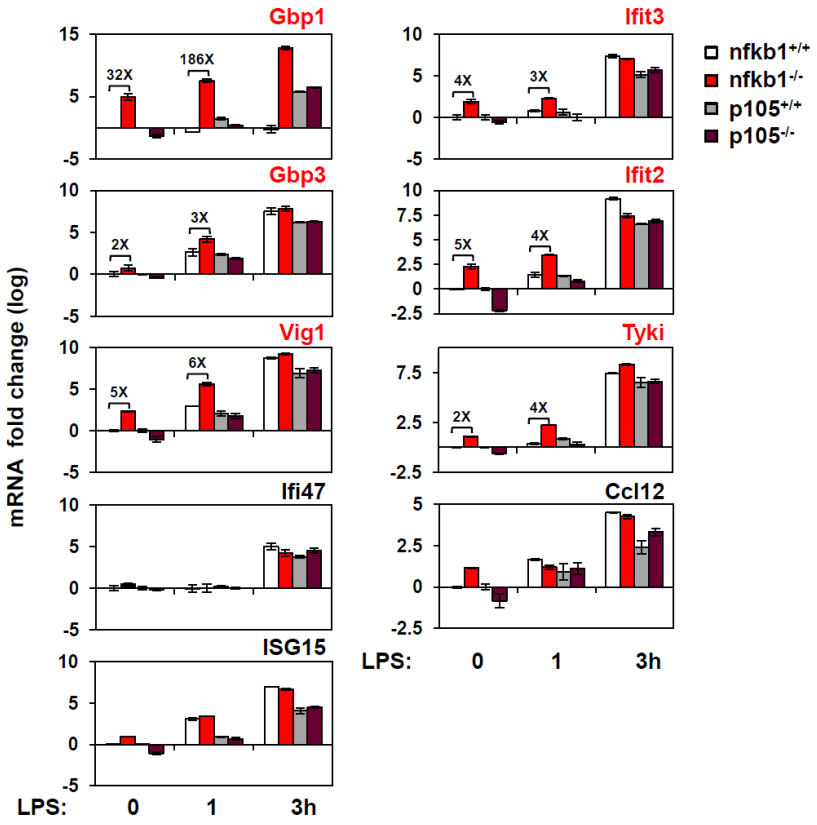
Figure 3.27 NFκB p50 promote proliferation in response to CpG

Percentage of the cell number change following wt (white), p50ko (red) and *ifnar*$^{-/-}$ (black) BMDM stimulation with LPS (0.1 µg/ml) or CpG (100 nM) relative to non-treated cells was determined by crystal violet assay in. Error bars represent the standard deviation from three biological replicates.



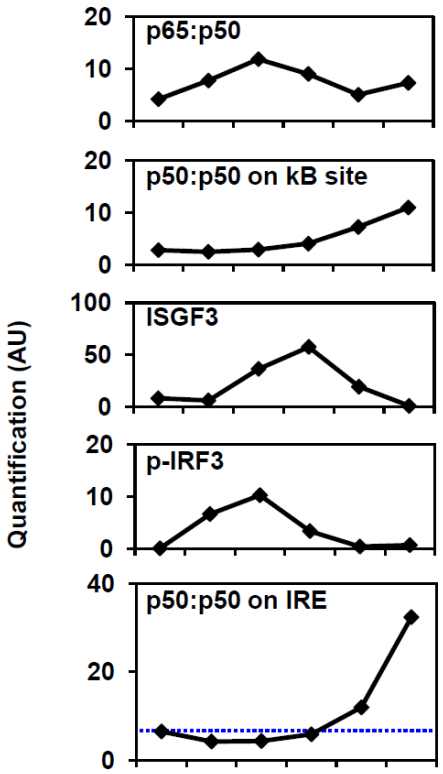
Appendix Figure 3.1 Expression profiling of IFNβ-responses reveals hyperexpression of some genes at early timepoints

Microarray mRNA expression profiles of wt and p50ko (*nfkb1*$^{-/-}$) BMDMs stimulated with IFNβ (10 U/ml) for 1 and 3 hour were matched to the K-means clusters in Fig.1D.



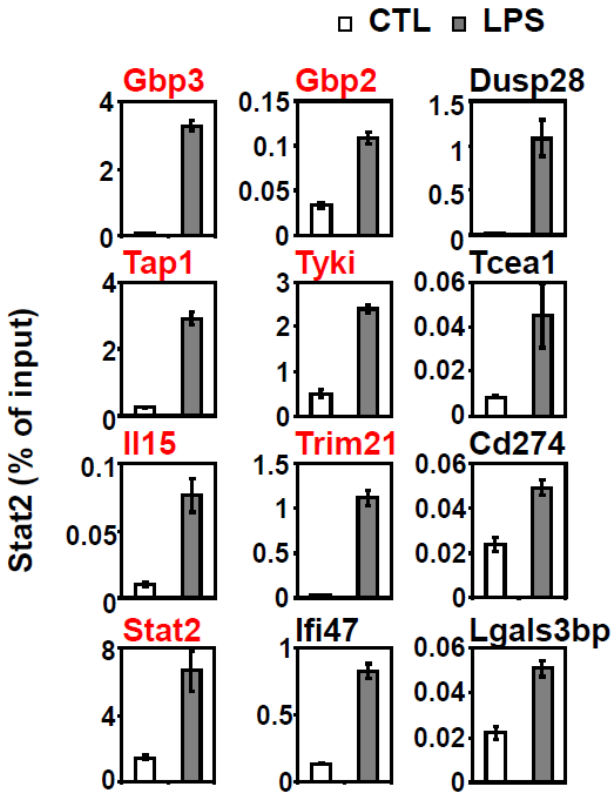
Appendix Figure 3.2 Gene expression phenotype observed in the p50ko cells is dependent on the absence of p50 rather than p105 protein

mRNA expression fold change of IFNβ inducible G-rich IRE genes (red) and non-G-rich IRE genes (black) were analyzed by RT-qPCR in nfkb1$^{+/+}$ (white), nfkb1$^{-/-}$ (red), p105$^{+/+}$ (grey), and p105$^{-/-}$ (purple) BMDMs stimulated with LPS (0.1 µg/ml) at 1 hour. Log expression fold change were plotted relative to each genotype's non-stimulated conditions.



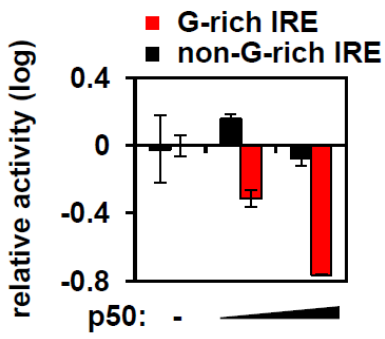
Appendix Figure 3.3 Transcription activator and p50:p50 repressor DNA binding activities in response to LPS

EMSA and nuclear western blots shown in Figure 1E were quantitated using ImageQuant software (GE Healthcare). κB-binding activities (p50:p65 and p50:p50) and IRE-binding activities (ISGF3 and p50:p50) were measured by EMSA. The activation profile of nuclear phospho-IRF3 (p-S396) was analyzed by Western blot.



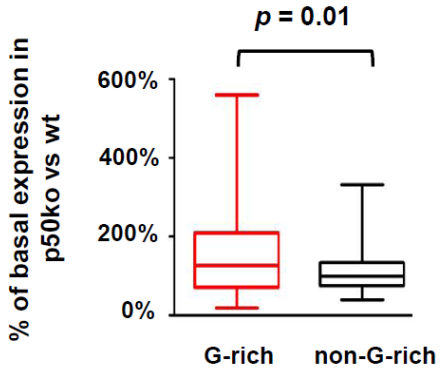
Appendix Figure 3.4 Stat2 is recruited to the IRE sequences whether they are G-rich or not

ChIP followed by qPCR analysis of the recruitment of Stat2 (ISGF3 complex) on p50-repressed, G-rich IREs (red) and p50-independent, non-G-rich IREs (colored in black) in response to LPS (0.1 µg/ml, grey bar) or PBS (CTL, white bar) for 1 or 2 hour in wt BMDMs.



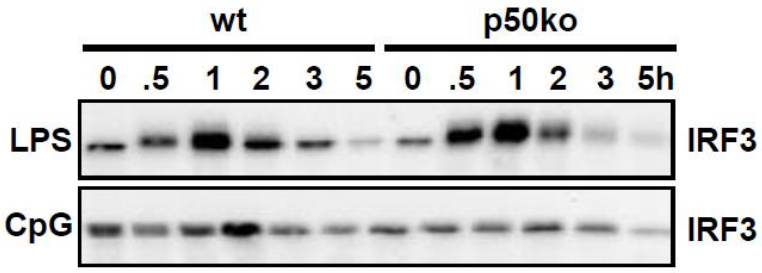
Appendix Figure 3.5 5 Increased expression of p50 protein reduced IRF-mediated activation of G-rich-IRE-driven reporter gene, but not non-G-rich-IRE driven reporter gene

TLR-HEK293T cells were transfected with IRE reporter plasmids (40ng) and p50 expression plasmid (0, 40 and 120 ng) for 48 hrs in 24 well plates. Cells were stimulated with LPS (0.1 ug/ml) for 6 hours. Relative IRE luciferase activity was graphed as a function of promoter activity in the absence of p50.



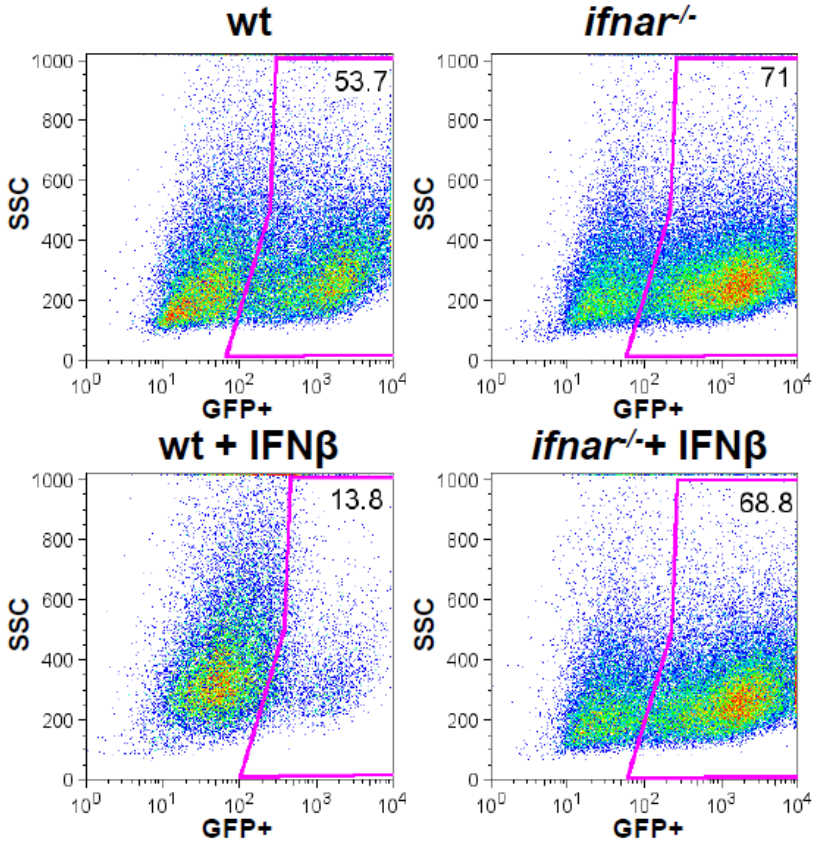
Appendix Figure 3.6 "G-rich" IREs are more likely to show elevated basal expression in p50-deficient BMDMs

Unpaired t-test was performed on % of basal expression in p50ko (*nfkb1*$^{-/-}$)vs wt BMDMs in duplicate microarray experiments, with G-rich (red) and non-G-rich IREs (black) from 107 interferon-β-inducible, IRE-containing genes. Edges of the boxplot indicate the 25th and 75th percentiles and whiskers indicate maximum and minimum data points.



Appendix Figure 3.7 IRF3 is not activated in response to CpG in both wt and p50ko BMDMs

The activation profiles for nuclear IRF3 were analyzed by immunoblots in response to LPS (0.1 ug/ml) or CpG (100 nM) using nuclear extract derived from wt and p50ko (*nfkb1*$^{-/-}$) BMDMs.



Appendix Figure 3.8 MCMV-GFP viral infection assay reveals IFNβ-mediated antiviral responses

A quantitative infection assay with MCMV-GFP, which reports only productive infections, to determine antiviral innate immune functions of macrophages. After a 24 hour mock treatment or treatment with IFNβ (100 U/ml), wt and *ifnar*$^{-/-}$ BMDMs were infected with MCMV-GFP. Virally infected cells expressing GFP were quantified (purple gate) by flow cytometry (side scatter over GFP plots) after 48 hours. Kristyn Feldman performed the anti-viral assays.

Chapter 4 : Edited miRNA seed signatures in IFN-mediated gene repression program

ABSTRACT

MicroRNA-mediated gene repression has emerged as a potent mechanism of biological regulation pertaining to cancer and immunity, yet identifying target genes of specific miRNAs has been difficult. We pursued a discovery strategy that is inclusive of post-transcriptional modifications to miRNA seed sequences. Starting with cell-type and stimulus-specific inflammatory gene repression programs, we identify the NF-κB-inducible miR-155 as a gene repression factor in fibroblasts and anti-inflammatory TH2 cells, but in macrophages and inflammatory TH1 cells a specific edited form of the miR-155 seed emerges as functionally relevant. In fact, edited forms of many highly expressed miRNAs emerge as candidate gene repression factors, in correlation with the strength of IFN signaling. Experimentally, we show that gene repression by specific edited forms of miRNAs is dependent on ADAR1, an interferon inducible, ISGF-3 controlled adenosine deaminase. We conclude that IFN-induced expression of ADAR1 controls inflammatory gene expression programs by mediating the widespread editing of miRNAs. Our work suggests that mapping the functional targets of miRNAs must consider cell type and stimulus-specific post-transcriptional miRNA modifications.

INTRODUCTION

Recent high throughput methods have uncovered a plethora of small RNAs in both animals and plants. For a few of these, mouse reverse genetic experiments have demonstrated a critical role in mammalian development and physiology (Stefani and Slack, 2008). One class of small RNAs are the non-coding ~22-nucleotide microRNAs (miRNAs) that induce post-transcriptional gene repression through imperfect base pairing with the 3'UTR region of the target mRNA. The 5' region of miRNA (particularly positions 2-7, known as the "seed") is the most conserved region of metazoan miRNAs, and has been shown experimentally and computationally to play a key role in target recognition (Doench and Sharp, 2004; Lai, 2002; Lewis et al., 2003). Functionally, this miRNA-messenger RNA (mRNA) interaction typically leads to post-transcriptional gene repression by inhibiting translation and by increasing mRNA degradation through an mRNA degradation pathway that may involve processing bodies (P-bodies) (Bagga et al., 2005; Jackson and Standart, 2007; Valencia-Sanchez et al., 2006). In fact, small reductions in the abundance of target mRNAs are often a reflection of larger changes in the protein abundance.

As miRNAs have emerged as physiologically important gene repression factors, there is pressing interest to identify the mRNA targets that are being regulated by them. A common strategy has been to match miRNA sequences, particularly the seed region, to potential target mRNAs using computational methods (John et al., 2004; Krek et al., 2005; Lewis et al., 2003). An inverse relationship between miRNA and target mRNA or protein levels may indicate direct regulation that is then tested using reporter genes that

include the 3'UTR sequence predicted to mediate regulation. The primary challenges of this approach are the short and imperfect base pair interactions between miRNA and mRNA and the relatively small changes seen in the mRNA abundance of many true mRNA targets. Furthermore, two assumptions of this approach are not well supported; it remains unclear whether miRNA expression levels correlate with their functional effect (given their role in RNA degradation pathways), and whether miRNA sequences always function as genetically encoded, or undergo a prior post-transcriptional modification such as an RNA editing step.

Indeed, the synthesis and degradation of miRNAs is tightly regulated through transcriptional synthesis, processing, maturation and half-life control. miRNAs were shown to undergo post-transcriptional modifications catalyzed by ADAR (adenosine deaminase acting on RNA) enzymes that result in A-to-I RNA editing of miRNAs (Blow et al., 2006; Habig et al., 2007; Kawahara et al., 2007a; Kawahara et al., 2007b; Nishikura, 2006). Editing events occur at specific sites over the pri-miRNA and pre-miRNA sequences and may control miRNA biogenesis (Yang et al., 2006) (Kawahara et al., 2007a), de-functionalize miRNAs and/or redirect miRNA target specificity when seed sequences are affected (Kawahara et al., 2007b).

Although statistical evaluation of potential target genes has been used to characterize the functional effects of specific miRNA seed sequences (Krutzfeldt et al., 2005; Lim et al., 2005), such an approach also offers a means to identify functional interactions without assumptions about specific sequence motifs. Indeed, statistical

methods have been used to identify *ab initio* sequences and candidate transcription factors regulating the promoters of specific gene expression modules (Segal et al., 2004). Post-transcriptional regulation of mRNAs was shown to affect the evolution of 3'-UTR sequences (Farh et al., 2005; Stark et al., 2005), suggesting that 3'-UTRs encode the susceptibility to regulatory control by miRNAs, and therefore allowing for an *ab initio* identification of regulatory motifs.

Inflammation is a complex response to pathogenic infection and tissue injury. It is tightly regulated, as misregulation of inflammation contributes to a variety diseases such as asthma, atherosclerosis, autoimmune disorders and cancer (Karin et al., 2006; Nathan, 2002). Inflammatory stimuli are known to activate hundreds of genes, but at the same time they also induce gene repression events that have not been well characterized. High throughput screening for miRNAs induced in response to inflammatory stimuli has only identified a few candidates (Landgraf et al., 2007; O'Connell et al., 2007; Taganov et al., 2006), with one, miR-155, playing important physiological roles by controlling B-cell development and proliferation as well as T- and dendritic cell function (Rodriguez et al., 2007; Thai et al., 2007). However, our understanding of how miR-155 mediates its physiological effects and which other miRNAs may be important remain incomplete.

Here, we took an alternate strategy to characterizing miRNA-mediated regulation of inflammatory responses. Beginning with the 3'UTRs of co-regulated genes in inflammatory gene repression programs, we employed bioinformatic and genetic approaches to identify enriched hexameric sequences. With this approach we not only

identified functional seed sequences of known miRNAs, but strikingly, uncovered that edited seed signatures dependent on ADAR1 play a major role in shaping inflammatory gene repression programs.

RESULTS

Inflammatory gene repression programs reveal the miR-155 seed signature

To characterize inflammatory gene repression, we analyzed publicly available microarray datasets of macrophages and B-cells treated with cytokines, pathogens, or pathogen-derived substances (Nau et al., 2002; Zhu et al., 2004). We found a substantial number of down-regulated genes, often more than up-regulated genes (Appendix Figure 4.1A and 4.1B). K-means clustering revealed that similar to gene activation programs, there is a core repression program common to all stimuli as well as stimulus-specific repression programs (Appendix Figure 4.1C). With our own microarray datasets from primary mouse embryonic fibroblasts (MEFs) and fetal liver-derived macrophages (FLDMs) stimulated with LPS (Supplementary Appendix Figure 4.1D) we found that the inflammatory response includes large gene repression programs that are not only stimulus-specific but also cell-type specific.

We then asked whether miRNAs may contribute to these gene repression programs. Computational analysis can reveal the statistical enrichment of a miRNA seed sequence, indicating its involvement in gene repression (Krutzfeldt et al., 2005). Similarly, we employed the Sanger 9.1 database, which contains 380 published miRNA sequences, and calculated seed count scores for the 3'UTR of each gene normalized to the length of the 3'UTR. Wilcoxon rank sum tests were performed to assess the statistical significance of the enrichment of each miRNA seed sequence in the 3'UTR of down-regulated genes by comparing seed count scores between down-regulated genes and genes without expression change. In addition to the miRNA seed sequences, we also

performed the same statistical test on all 4,096 possible hexamer RNA motifs in order to compare the statistical significance of the miRNA seed enrichment.

To validate this computational method and its reliance on the seed sequence, we performed the seed enrichment analysis on a microarray dataset from B-cells over-expressing miR-155 (Costinean et al., 2006), one of the very few miRNAs highly induced in response to inflammatory stimuli. Indeed, the miR-155 seed sequence was found to be highly over-represented among all 4,096 possible hexamer motifs in the 3'UTR sequences of genes down-regulated upon transgenic overexpression of miR-155 in B-cells (Figure 4.1A). This finding not only validated our approach but also suggests that miR-155 in particular does affect widespread mRNA transcript decay, and does so by base pairing with the 3'UTR of target mRNAs through the 2-7 seed sequence.

We further carried out the seed enrichment analysis on down-regulated genes from microarray studies in MEFs, B-cells (Zhu et al., 2004) and macrophages upon LPS stimulation. We found that the miR-155 seed was indeed over-represented (top 5% among all 4,096 hexamers) in the 3'UTR of LPS down-regulated genes in MEFs and B-cells (Figure 4.1B), suggesting that induced expression of this miRNA may be involved in the regulation of gene repression programs. However, to our surprise we did not find it over-represented in LPS-repressed genes in macrophages (Figure 4.1B). Northern blotting showed that miR-155 is induced by LPS in MEFs in an NFκB-dependent manner (using *rela*$^{-/-}$*crel*$^{-/-}$*nfkb1*$^{-/-}$ MEFs lacking the three canonical NFκB/Rel proteins) and we found strong inducible expression not only in B-cells but also in macrophages (Figure

4.1C). These findings provide evidence that miR-155 is functional in down regulating genes in MEFs and B-cells, but, despite robust expression, no such functional evidence was found in macrophages.

IFN activated ADAR-1L may be responsible for miR-155 A-to-I editing

Given the lack of miR-155 seed enrichment in macrophages, we considered that RNA editing may de-functionalize miR-155, as had been shown for pri-miR-142 in hematopoietic tissues (Yang et al., 2006). ADAR1 is an adenosine deaminase shown to mediate A-to-I RNA editing of miRNAs (Kawahara et al., 2007a; Kawahara et al., 2007b; Nishikura, 2006; Yang et al., 2006), and one of its isoforms, ADAR-1L (150 kDa protein), is inducible by a number of inflammatory agents (George and Samuel, 1999; Patterson and Samuel, 1995; Yang et al., 2003). Quantitative RT-PCR showed that LPS-induced expression of ADAR-1L is much stronger in macrophages than in MEFs (Figure 4.2A). The induced expression of ADAR-1L is dependent on the type I IFN pathway as the induction in response to LPS is completely abolished in $IFNAR^{-/-}$ macrophages. Furthermore, ADAR-1L induction in response to IFNβ stimulation is intact in $NF\kappa B^{-/-}$ (*rela*$^{-/-}$*crel*$^{-/-}$*nfkb1*$^{-/-}$) cells (Figure 4.2B). We identified a putative ISRE (interferon stimulatory response element) in the mouse ADAR-1L promoter (position -55 to -66). To determine whether ADAR-1L is a direct target gene of the IFN pathway, we undertook chromatin immunoprecipitation (ChIP) for the IFN-induced transcription factor ISGF3 components Stat1 and Stat2. These studies revealed the LPS-responsive recruitment of Stat1 and Stat2 to the ADAR-1L promoter, while the NFκB/RelA target promoter of the IκBα gene did not show Stat1/2 association (Figure 4.2C).

Our data suggested that cells that undergo strong IFN signaling may de-functionalize the miRNAs by inducing the expression of ADAR1L. However, we also considered the possibility that miRNA editing may occur in seed sequences, redirecting their target specificity, as I:C and G:C base pairs make equivalent contributions to the hybridization between miRNAs and their targets (Kawahara et al., 2007b). With two adenines present in the miR-155 seed sequence, RNA editing by ADAR-1L could potentially result in three additional seed sequence specificities (Figure 4.2D). We termed the unedited or native form miR-155_N, and the edited forms _E1, _E2, and _E3, and tested them in the Rank-sum test seed enrichment analysis. Interestingly, the seed sequence of one of the edited forms, miR-155_E1, was indeed enriched in down-regulated genes in macrophages (but not in MEFs), but neither of the other two edited sequences were found to be enriched in any dataset (Figure 4.3A and 4.3C and Appendix Table 4.1). Together these observations suggest the hypothesis that stimuli activating the type I interferon pathway reprogram the specificity of the miR-155 gene repression factor *via* A-to-I RNA editing of a specific adenine in its seed sequence.

We first examined this hypothesis by ascertaining the specificity of the bioinformatic analysis. We obtained littermate MEFs and peritoneal macrophages (PMs) from wild-type and bic-/miR-155-deficient mice, treated them with LPS and analyzed the transcriptome using microarrays. We found that LPS stimulation led to the previously observed cell-type-specific gene repression programs, which we characterized by k-means clustering. Seed enrichment analysis of the 3'UTRs of down-regulated genes

confirmed the statistical enrichment of the native form of the miR-155 seed sequence (N) in wild type MEFs but not in the knockout counterparts (Figure 4.3C and 4.3D and Appendix Table 4.1). Importantly, the miR-155_E1 was again over-represented in the dataset from wild type PMs but not in bic/miR-155-deficient PMs stimulated with LPS (Figure 4.3A and 4.3B and Appendix Table 4.1).

We further confirmed the functional specificity of edited and native miR-155 seed sequences with quantitative RT-PCR of individual target genes. Macrophage metalloelastase (MMP-12), a pro-inflammatory mediator associated with various inflammatory diseases, such as atherosclerosis, aneurysms, inflammatory skin and pulmonary diseases and cancers (Nenan et al., 2005), contains one edited miR-155 but no native miR-155 seed sequence in its 3'-UTR region. LPS induced down regulation of MMP-12; however, this gene repression event is much attenuated in miR-155 knockout macrophages but not in bic/miR-155-deficient MEFs (Figure 4.4A). Similarly, the LPS-induced down regulation of a cytochrome gene Cyp51, which only has an edited miR-155 seed sequence in its 3'-UTR, is attenuated in the bic/miR-155-deficient macrophages but not in MEFs (Figure 4.4B). In contrast, down regulation of native miR-155 targets, such as transcription factor AP-2β (Tcfap2ab) and zinc finger homeodomain 4 (Zfhx4), were attenuated in the bic/miR-155-deficient MEFs (Figure 4.4C and 4.4D). Together, our results suggest a functional role of the edited miR-155 seed sequence in mediating inflammatory gene repression program in activated innate immune cells.

Enrichment of the edited miR-155 seed signature in pro-inflammatory adaptive immune cells

Next, we examined whether edited or unedited miR-155 seed sequences were enriched in gene repression programs induced in adaptive immune cells, such as B- and T-cells. Interestingly, in unactivated pre-B-cells that expressed miR-155 ectopically from a transgene (where the expression of miR-155 is elevated but IFN signaling and ADAR-1L expression are not activated) only the native miR-155 seed sequence was highly enriched in repressed genes (Figure 4.1A and 4.5A). On the contrary, in activated B-cells stimulated with LPS (Zhu et al., 2004), where ADAR-1L expression is activated, both the edited miR-155_E1 and the native miR-155 seed sequences were highly enriched in the 3'UTR of down-regulated genes (Figure 4.5B). Examining the 3'-UTR of genes misregulated in bic/miR-155-deficient splenic B-cells (Rodriguez et al., 2007), we found that both the native miR-155 and edited E-1 seed sequences were enriched (Appendix Figure 4.2). Focusing on miR-155 function in T-cells, TH1 and TH2 cells were distinguished (Rodriguez et al., 2007). In the 3'UTR of genes hyper-expressed in the bic/miR-155 knockout TH1 cells, both the edited miR-155_E1 and the native seed sequences were enriched (Figure 4.5D). On the contrary, only the native miR-155 seed sequences were enriched in the analogous dataset of TH2 cells (Figure 4.5C). The same trend of edited and native miR-155 seed enrichment was observed with a variation of seed sequence selections (Appendix Table 4.2). These results support the hypothesis that miR-155 seed editing that alters its mRNA target specificity is associated in a variety of cell types with inflammatory signaling.

IFN and ADAR-1L dependent enrichment of many edited miRNA seeds

To determine whether miRNA editing could be a more widespread phenomenon affecting many miRNAs, we generated a library of hexamer sequences of all potential edited miRNA seed sequences from a total of 380 miRNAs downloaded from the Sanger9.1 database. The 4096 possible hexamer RNA motifs are thus partitioned into three classes: miRNA native seed sequences (290), possible edited miRNA seeds (628), and the remaining hexamers (3174) that may not have biological relevance. The edited miRNAs were termed according to the classification system we first developed for miR-155 derived forms (Figure 4.2D). Surprisingly, we observed a large number of edited miRNA seed sequences enriched in the 3'UTR of LPS down-regulated genes in macrophages, whereas the pool of native miRNA seed sequences was not enriched (Figure 4.6A). Interestingly, this enrichment of edited miRNA seeds was abolished in $IFNAR^{-/-}$ macrophages, with native miRNA seed sequences rising to the top of the list (Figure 4.6B). Similarly, correlating with a much weaker type I IFN pathway and lower ADAR-1L expression, enrichment of edited miRNA seeds was not observed in the MEF dataset (Figure 4.6C) or in the TNF-induced gene repression program in macrophages (Figure 4.6D).

We examined which miRNAs emerged in our analysis as particularly strong candidate gene repression factors. Interestingly, an edited form of 10 out of 11 known inflammation-inducible miRNAs (in which the seed sequences can be edited) is enriched in the LPS-repressed macrophage data set (Figure 4.7A and Appendix Figure 4.3). In addition, some of the most highly enriched hexamers in LPS-induced macrophages

matched the edited seed sequences of some miRNAs that are constitutively expressed at particularly high levels in Raw cells, a macrophage cell line (Figure 5.7B and Appendix Figure 4.3)(Liu et al., 2004). These include specific edited forms of let-7, miR-130, miR-125, miR-214, miR-146 and miR-30a. Particularly, an edited form of miR-376, which was previously shown to redirect target specificity (Kawahara et al., 2007b), appeared to be one of the most highly enriched hexamers (Appendix Figure 4.3).

To characterize the role of ADAR-1L on the expression of endogenous miRNA target genes, we performed siRNA knockdown of ADAR-1L in peritoneal macrophages. LPS induced expression of ADAR-1L was reduced up to 3 fold by the ADAR-1L knockdown as compared to the non-targeting siRNA knockdown control (Appendix Figure 4.4). Among many of the potential targets of edited miRNAs (Appendix Table 4.3), we specifically selected transcription factors Cebp-α and Jun for their functional importance in gene regulation and the enrichment of edited seed sequences in their 3'-UTR region (Figure 4.8A). Quantitative RT-PCR confirmed that these genes are down-regulated in response to LPS. However, when knocking down ADAR-1L, LPS-induced gene repression of Cebp-α and Jun was abolished (Figure 4.8B). Similarly, ADAR-1L knockdown also led to hyper-expression of the dolichyl pyrophosphate phosphatase (Dolpp1), which has enriched edited seed sequences in its 3'-UTR and shows only mild changes in mRNA abundance in unaltered macrophages. Among the many edited miRNA seed sequences in these ADAR-1L target genes, there are four edited let-7 seed sequences (let7_E2) in the 3'-UTR of Cebp-α and three edited miR-125 seed sequences (miR-125_E1) in Dolpp1. Together, our results suggest IFN-induced ADAR1 may be

responsible for global miRNA editing that causes a stimulus-responsive switch in target specificity of miRNAs, and thereby induces gene repression programs, even when miRNA expression itself may not be induced.

DISCUSSION

Stimulus-induced gene repression was observed in the earliest microarray studies (Iyer et al., 1999) but has not been widely characterized. Gene repression may involve the inhibition of transcriptional initiation but mRNA half-life control is also a critical determinant. miRNAs may accelerate mRNA degradation, either as the primary means of affecting gene repression, or as a secondary effect of inhibiting translation. Indeed, we find that the seed sequences of specific miRNAs are statistically over-represented in down-regulated mRNAs and we have confirmed the specificity of the statistical method by comparing the stimulus-induced gene repression programs in wild type and miR-155 deficient cells.

Though the role of tissue-specific expression of miRNAs in developmental processes is well documented (Calin and Croce, 2006; Stefani and Slack, 2008; Wienholds and Plasterk, 2005), studies of miRNA function in genome-wide stimulus-responsive expression programs have been lagging, despite reports of a few stimulus-inducible miRNAs (O'Connell et al., 2007; Pedersen et al., 2007; Taganov et al., 2006). In this first attempt to characterize the potential functional effect of miRNAs in mediating stimulus-induced gene repression programs, we chose miR-155 as a model system because its functional importance in immune cells was established (Costinean et al., 2006; Rodriguez et al., 2007; Thai et al., 2007) and its expression is highly induced in response to a variety of inflammatory stimuli (Dahlberg and Lund, 2007; O'Connell et al., 2007; Rodriguez et al., 2007; Taganov et al., 2006; Thai et al., 2007). A discrepancy

between miR-155 expression in macrophages and a lack of its functional signature from native miR-155 seed sequence led us to a new classification scheme of edited miRNA seed sequences, and the bioinformatic identification of a specific edited miR-155 form. Indeed, even in our global analysis of miRNA-mediated repression, we consistently find only one of several possible edited miR-155 seed sequences to be functionally relevant in gene repression. Our interpretation of these functional phenotyping studies was supported by genetic evidence using miR-155 knockouts.

Inducible A-to-I editing activity was observed in spleen, thymus and peripheral lymphocytes from endotoxin-treated mice and ConA/IL-2 stimulated splenocytes (Yang et al., 2003). Particularly, ADAR1 has been shown to be highly inducible by a number of inflammatory agents (George and Samuel, 1999; Yang et al., 2003). Our results demonstrating that type I IFN/ISGF3-dependent induction of ADAR1 is much stronger in macrophages stimulated with LPS than with TNF (or in non-immune cells such as fibroblasts) suggests that ADAR1 may function to reprogram the target specificity of miR-155 in the context of immune responses that involve interferons. This is not restricted to innate immune cells (exemplified by macrophages), but also pertains to adaptive immune cells such as pro-inflammatory TH-1 T-cells which reveal the edited miR-155 seed signature in their miR-155 regulated transcriptome, in contrast to anti-inflammatory TH-2 T cells which reveal the native miR-155 form.

Our discovery pipeline was based on identifying functional seed sequence signatures but circumvented the assumptions that genetically encoded seed sequences are

functional (allowing us to identify edited seeds) and that expression levels of miRNA are correlated with their functional effects. Indeed, editing of miRNA may lead to degradation by Tudor-SN, a component of RISC and a ribonuclease specific for inosine-containing dsRNAs (Yang et al., 2006). The edited miR-142 was cloned after treating cells with 2'-deoxythimidine 3', 5'-biphosphate (pdTp) (Yang et al., 2006), a specific inhibitor of Tudor-SN and I-dsRNA-specific ribonuclease that prevents rapid degradation of inosine-containing dsRNAs (Caudy et al., 2003; Scadden, 2005). However, such inhibition skews quantitative estimates of edited miRNA abundances and may reveal a variety of edited forms that are not functionally relevant for gene repression. The fact that miRNAs may be edited has also been shown using cloning strategies (Blow et al., 2006; Kawahara et al., 2007a; Kawahara et al., 2007b; Yang et al., 2006), though ADAR1L-mediated editing may to-date be under-reported because it is only highly expressed following inflammatory exposure and functions in the cytoplasm, rather than on immature nuclear miRNA forms. Our work demonstrates that seed editing has profound functional consequences on immune gene repression programs, and thus should motivate further large scale unbiased sequencing of miRNAs in a variety of immune-related physiological contexts. We speculate that the editing reaction may in some sense "functionalize" miRNAs increasing their specific activity, while at the same time possibly reducing their half-life.

Our results suggest that miRNA editing does not only occur in the miR-155 seed sequence, but may pertain to many other miRNAs as well. Thereby, innate immune gene repression programs that are induced in response to pathogen infection, may not

necessarily involve the induced expression of miRNAs but merely the retargeting of constitutively expressed miRNAs. Editing of mature miR-376 for example, has been shown to occur at the seed sequences (Kawahara et al., 2007b), and we have observed edited miR-376 seed sequences to be one of the highly enriched signature of inflammatory repression program in wild-type but not IFNAR$^{-/-}$ macrophages. Genetic deletion of ADAR1 causes embryonic lethality due to widespread apoptosis in the embryonic liver (Hartner et al., 2004; Wang et al., 2004), indicating a major role in hematopoiesis (including myeloid differentiation into macrophages) and immune function. Indeed, while our bioinformatic analysis utilized gene repression programs (and the mRNA half-life control implicit in these), it is likely that miRNA editing will also affect induced gene expression programs. Our finding that many of the IFN-inducible anti-viral miRNAs (Pedersen et al., 2007) leave their functional signature in down-regulated mRNAs as specific edited seeds rather than native seeds (Fig. 6e), may also suggest a potential anti-viral function for the widespread editing catalyzed by interferon-inducible ADAR-1L. More generally, our work demonstrates that cell type-specific and stimulus-specific post-transcriptional modifications of miRNAs can be determinants of large gene expression programs.

MATERIAL AND METHODS

Cell Culture and Reagents.

Primary mouse embryonic fibroblasts (MEFs) were generated from C57/BL6 (wild-type) and *crel*$^{-/-}$*p65*$^{-/-}$*nfκb1*$^{-/-}$ (termed NF-κB$^{-/-}$) E12.5-14.5 embryos and used

between passage 4-7. bic/miR-155$^{-/-}$ and wild-type littermate MEFs (Thai et al., 2007) were used in microarray experiments. Confluent cells were serum starved (0.5% serum) for 24 hours before stimulated with either 0.1 μg/ml LPS (Sigma, B5:055) or 2500 U/ml IFNβ (IFNβ was a gift from Biogen, Inc). Peritoneal macrophages (PMs) were isolated by peritoneal lavage 3 days following peritoneal injection of 2.5 ml 3% thioglycollate. PMs were isolated from bic/miR-155$^{-/-}$ and bic/miR-155$^{+/+}$ mice on the same mixed 129SV X C57BL/6 genetic background for microarray experiments. Cells were plated in RPMI medium 1640 with 10% fetal bovine serum (FCS) for 5 hours and then starved (0.5% serum) for overnight before stimulated with 0.1 μg/ml LPS. Bone marrow derived macrophages (BMDMs) were generated by culturing bone marrow progenitors in L929 cell-conditioned medium (L929CM) for 7 days and then stimulated with 0.1 μg/ml LPS. Splenocytes were isolated from spleens by dissociation through sterile 40-μm cell strainers (BD Falcon). Red blood cells were lysed with red blood cell lysis buffer (eBioscience). Splenocytes were resuspended in 5% FCS and stimulated with 40 μg/ml LPS.

microRNA Northern Blot.

RNA was extracted using Trizol Reagent (Invitrogen) and the detection of miR-155 was carried out by polyacrylamide gel electrophoresis Northern methods as previously described (Bagga et al., 2005).

Messenger RNA Detection.

Total RNA was isolated using RNeasy mini kit (Qiagen) and converted to cDNA by using SuperScript II reverse transcriptase (Invitrogen) both according to manufacturer's protocol. Resulting cDNA was quantified with real-time quantitative PCR (SYBRgreen) analysis with specific primer pairs for each mRNA transcript. The primer sequences used in the qPCR analysis were: 5'-TTCAGGGGACCCACAGG-3' (ADAR-1L.f), 3'-GCGGGTATCTCCACTTGCTA-3' (ADAR-1L.r), 5'-AACTTTGGCATTGTGGAAGG-3' (GAPDH.f), 5'-GGATGCAGGGATGATGTTCT-3' (GAPDH.r).

ADAR1L knockdown.

Peritoneal macrophages were transfected 12 hrs after plating with a siGenome Smart pool siRNA reagent against murine ADAR1 (Dharmacon NM_019655) using DeliverX Plus transfection reagent (Panomics). LPS-stimulations were begun 12 hrs after transfections.

Chromatin Immunoprecipitation (ChIP) Assay.

ChIP assay was performed as previously described (Perissi et al., 2004). Immunoprecipitated DNA fragment was quantified by real-time quantitative PCR analysis. Anti-Stat1 and anti-Stat2 antibodies used in immunoprecipitation experiments were from Santa Cruz Biotechnology. The primer sequences used in the qPCR analysis were: 5'-AAACCCCTCCCTCCTCTTG-3' (ADAR-1L_ISRE.f), 5'-CACCTGTGGCCGTAAGATG-3' (ADAR-1L_ISRE.r), 5'-

GCTTCTCAGTGGAGGACGAG-3' (IκBα_κB.f), CTGGCAGGGGATTTCTCAG-3' (IκBα_κB.r).

Microarray Experiment.

Total RNA was isolated using RNeasy Mini kit (Qiagen). Sample preparation and hybridization to Codelink mouse Uniset 1 (GE Healthcare) microarrays and Illumina Mouse Refseq v1.1 Expression BeadsChIP microarrays (20k, BD-26-213) were performed at the UCSD Biogem facility.

Microarray analysis.

For figure 1a, 1b and 4a, genes down-regulated by overexpression of a miR-155 transgene in pre B-cells were obtained from table 3 of Costinean et al. (Costinean et al., 2006) for seed enrichment analysis. For figure 1b, data was obtained from microarray experiments using MEFs and PMs and GE Codelink Uniset 1 microarrays. For figure 3a and 3b, microarray experiments with wt and bic/miR-155$^{-/-}$ thioglycollate-elicited peritoneal macrophages stimulated with LPS at 0 and 24 hours were performed with Illumina Mouse Refseq v1.1 microarrays; down-regulated genes were selected for seed enrichment analysis using a 2-fold cut-off. For figure 3c and 3d, microarray experiments with bic/miR-155$^{-/-}$ and wild-type littermate MEFs stimulated with 0.1 μg/ml LPS at 0 and 24 hours were performed with Illumina Mouse Refseq v1.1 microarrays; down-regulated genes were selected for seed enrichment analysis using a 1.68-fold cut-off. For figure 4b, down-regulated genes were selected from the 2937 differentially expressed features from Table III of Zhu et al. (Zhu et al., 2004) for seed enrichment analysis as

genes with at least 2-fold down regulation. For figure 4c and 4d, microarray data were obtained from Rodriguez et al. (Rodriguez et al., 2007). Genes that are differentially up regulated in bic/miR-155$^{-/-}$ TH1 and TH2 cells as compared to bic/miR-155$^{+/+}$ cells were selected by SAM analysis for further seed enrichment analysis. For figure 5a, 5b, 5c, 5d and supplementary figure 3, microarray experiments of wild-type and IFNAR$^{-/-}$ thioglycollate-elicited peritoneal macrophages stimulated with 0.1 μg/ml LPS or 1 ng/ml TNF at 0 and 6 hours were performed with GE Codelink mouse Uniset 1 microarrays, down-regulated genes were selected for seed enrichment analysis using a 2-fold cut-off. For supplementary figure 1a, publicly available microarray data for human macrophages were obtained from Nau et al. (Nau et al., 2002) 977 significantly changed genes with at least 3-fold change on exposure to one or more of the pathogens (E. Coli, EHEC, S. typhi, S. tyhirnurium, S. aureus, L. monocytogenes, M. tuberculosis and BCG) at one or more of the 1, 2, 6, 12 and 24 hours time points as described in the paper were selected for K-means clustering analysis with correlation uncentered similarity matrix (Eisen et al., 1998). For supplementary figure 1b, microarray data of 2937 differentially expressed features upon a number of inflammatory stimulation in B-cells were obtained from Zhu et al. (Zhu et al., 2004) and clustered using K-means clustering analysis. For supplementary figure 1c, microarray experiments of wild-type MEFs stimulated with 0.1 μg/ml LPS (Sigma), 1 ng/ml murine TNF (Roche), or 1 ng/ml IL-1 (Calbiochem) at 0, 1 and 8 hours were performed with Codelink mouse Uniset 1 microarrays. Differentially induced and repressed genes were selected for K-means clustering analysis as genes with at least 2 fold change in one or more stimuli or time point conditions. For supplementary figure 1d, microarray experiments of wild-type MEFs and fetal liver-derived macrophages

(FLDMs) stimulated with 0.1 µg/ml LPS at 0, 1, 3 and 8 hours were performed with Illumina Mouse Refseq v1.1 Expression BeadsChIP microarrays. Differentially induced and repressed genes were selected for K-means clustering analysis as genes with at least 1.68 fold change in one or more stimuli or time point conditions. For supplementary figure 2, the list of upregulated genes from miR-155-deficient B-cells as compared to wt B-cells stimulated with LPS and IL-4 for 24 hr were obtained from Vigorito et al. (Rodriguez et al., 2007) Upregulated genes were compared against the total genes on the Affymetrix 430 2.0 Genechip arrays for seed enrichment analysis.

Potential edited miRNA target genes

For supplementary table 3, edited miRNA seed sequences enriched in the 3'-UTR of LPS repressed genes in wild-type but not IFNAR$^{-/-}$ PMs were selected. Genes that are more down-regulated in wild-type than in IFNAR$^{-/-}$ PMs and have more than 10 edited seed sequences in the 3'-UTR were selected as potential targets. Edited seed counts represent the total number of edited seed sequences (derived from the Sanger9.1 database) present in the 3'-UTR of the gene. The total numbers of edited seed sequences that are derived from the subset of miRNAs that have been validated by traditional cloning methods and/or Northern are indicated in the last column.

Seed Enrichment Analysis.

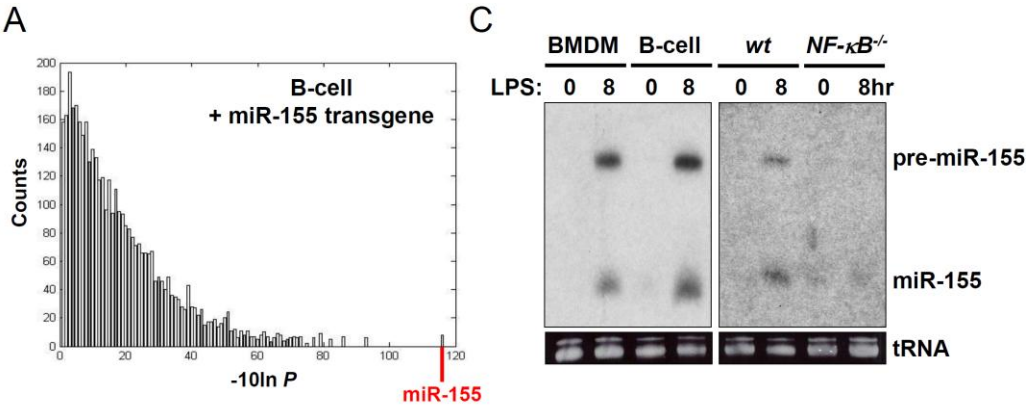
For each of the 4,096 possible hexamer or 16,384 possible 7mer RNA motifs, the number of occurrences in the 3'UTR of each mRNA transcript divided by the length of the 3'UTR was recorded as seed count scores. 3'UTR sequences were downloaded from

UCSC Genome Bioinformatics (mouse mm8 build). A non-parametric test (one-tailed Wilcoxon rank sum test) was performed on the distributions of seed count scores between targeted mRNA transcripts (For example, LPS down-regulated genes) and transcripts without expression change (or total transcripts on the microarray) for each of 4,096 or 16,384 RNA motifs. A *P* value was obtained for each of the RNA motifs and the distributions of the negative natural logarithm of all 4,096 or 16,384 *P* values were graphed. Percentile of rank position was defined from the top of the list of all 4,096 or 16,384 RNA motifs. miRNA sequences were downloaded from Sanger 9.1 database (miRBase). 227 experimentally validated miRNAs were defined as miRNAs that have been cloned or validated with Northern blot analysis. Edited seed sequences were generated by converting nucleotide sequences from A to G in all possible combinations for each miRNA seed sequence. When plotting *P* value distributions in stacked bar graphs, if a non-edited seed is the same sequence as an edited seed, this RNA motif will be graphed as non-edited seed. Non-seed hexamers in stacked bar graphs were defined as the set of the 4,096 RNA motifs that did not match non-edited and edited seeds. A detailed description of the microarray data and analysis is presented in the supplementary methods.

ACKNOWLEDGEMENTS

I thank Dr. To-Ha Thai and Dr. Klaus Rajewsky for the mir-155$^{-/-}$ mice, Janette Holtz for the miRNA northern blot, Dr. Amy Pasquinelli for critical suggestions and Dr. Christopher Glass for letting me use their microarray data. I thank Dr. N. Rajewsky and

Dr. C. Elkan for advice on statistical methods, P. Loriaux for critical reading and Dr. G. Gosh, and members of the Signaling Systems Laboratory for discussion and support. C.S. Cheng was supported by Department of Defense Breast Cancer Research Program Predoctoral Traineeship Award.



B

Cell Type	Percentile of 290 miRNAs seeds	Percentile of 4,096 hexamers
MEF + LPS	2%	3%
B-cell + LPS	7%	5%
MAC + LPS	38%	44%

Figure 4.1 miR-155 seed is enriched in the 3'UTR of repressed genes

(A) Wilcoxon rank sum test analysis revealing the enrichment of the native miR-155 seed sequence (relative to all 4096 possible hexamer sequences) in the 3'UTR of genes repressed in pre-B-cells expressing a miR-155 transgene (Zhu et al., 2004). The number of hexamers scoring within specific p-value bins are graphed. The position of miR-155 is indicated in red. (B) The rank position percentile of the native miR-155 seed sequence determined by the Wilcoxon seed enrichment analysis in the gene repression programs of LPS-stimulated MEFs, B-cells (Zhu et al., 2004) and peritoneal macrophages (MAC). The percentile (from the top) of the miR-155 seed is shown relative to the seed sequences of 290 unique miRNA seed sequences from the Sanger 9.1 database, and relative to 4,096 possible hexamer RNA motifs. (C) Northern Blot for miR-155 showing the expression of the 70 nt precursor and the 22 nt mature form in bone marrow derived macrophages (BMDM), B-cells, wild-type fibroblasts (wt) and NF-κB-deficient fibroblasts (NF-κB$^{-/-}$) in response to 8 hr stimulation with LPS (40 μg/ml for B-cells and 0.1 μg/ml for BMDM and fibroblasts).

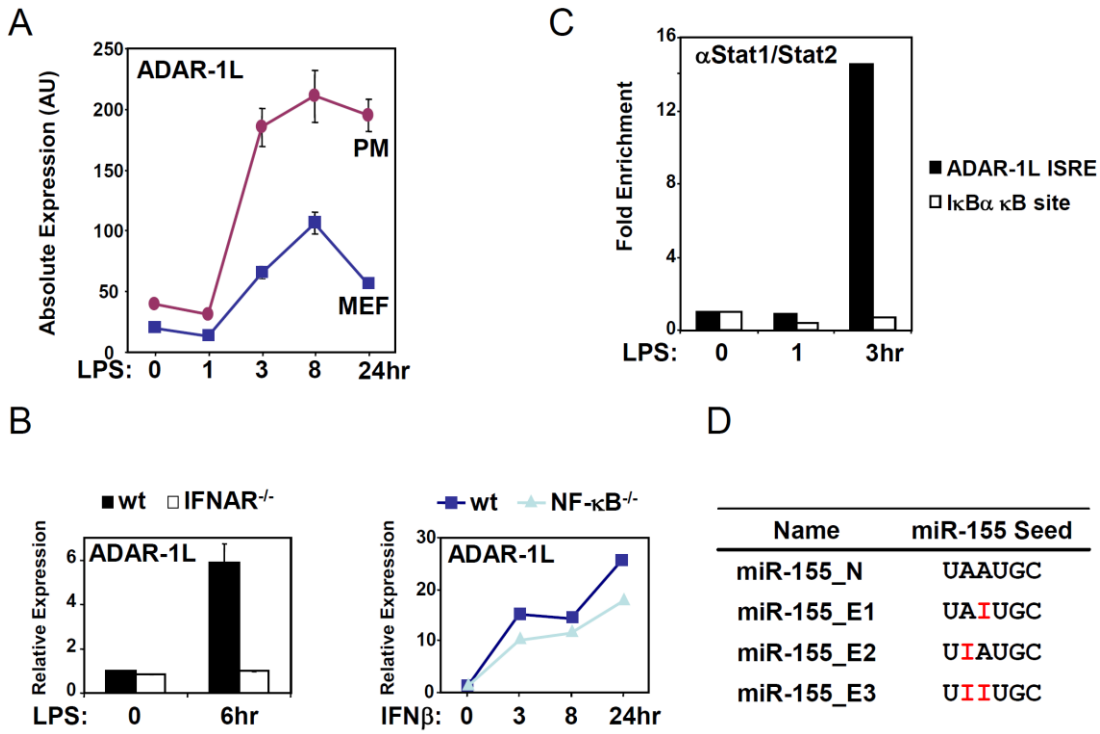


Figure 4.2 Type I IFN activated ADAR-1L may be responsible for miR-155 A-to-I editing

(A) ADAR-1L mRNA expression in peritoneal macrophages (PMs, red)) and MEFs (dark blue) after LPS (0.1 µg/ml) stimulation for 0, 1, 3, 8, 24 hours were analyzed by qPCR following reverse transcription (RT-qPCR). (B) Left, ADAR-1L mRNA expression in wild-type (black) and IFNAR$^{-/-}$ (white) PMs stimulated with LPS (0.1 µg/ml) for 0 and 6 hours, as revealed by microarray studies. Right, ADAR-1L expression in wild-type (dark blue) and NF-κB$^{-/-}$ (light blue) MEFs stimulated with LPS (0.1 µg/ml) for 0, 3, 8 and 24 hours, as revealed by RT-qPCR. (C) Occupancy of the ISGF3 complex on the consensus ISRE site in the promoter of ADAR-1L was revealed by ChIP for Stat1/2 followed by qPCR (black). The absence of ISGF3 recruitment on the κB site in the promoter of IκBα (white) serves as a negative control.(D) Three possible A-to-I edited seed sequences of miR-155 are represented as miR-155_E1, miR-155_E2 and miR-155_E3. The original unedited miR-155 seed sequence is designated miR-155_N (for nascent or native).

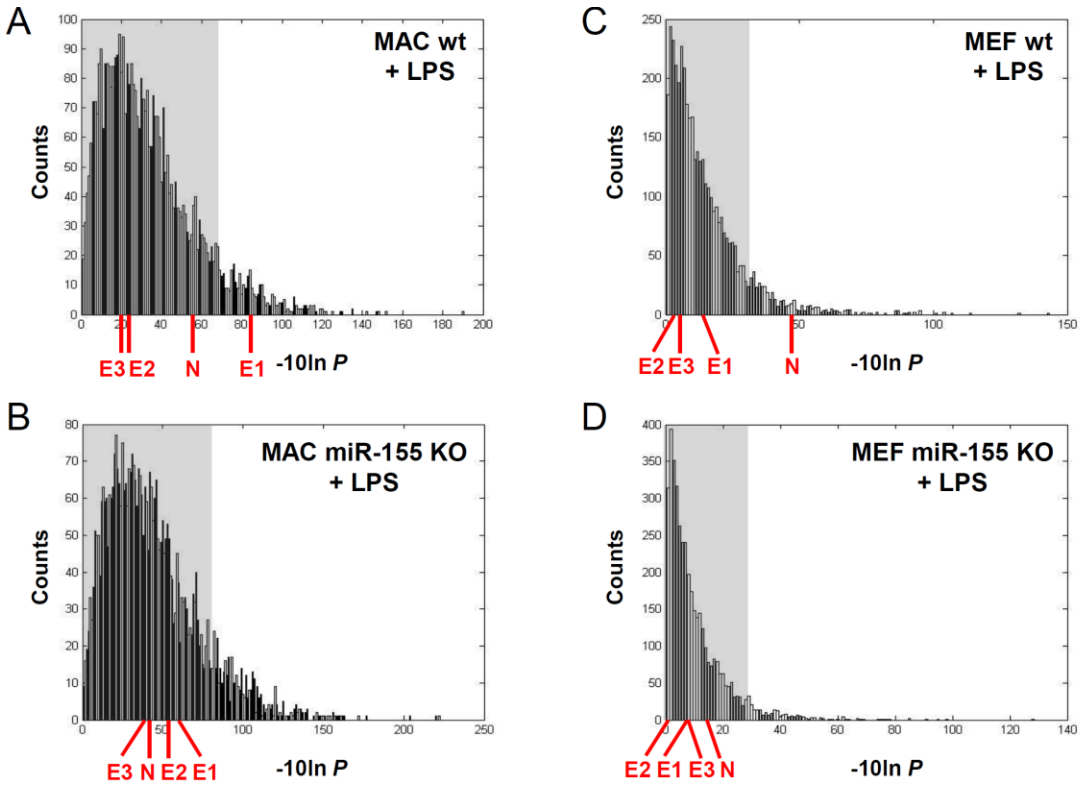


Figure 4.3 Gene repression by an edited miR-155 in activated innate immune cells
(A) Seed enrichment analysis was performed on LPS down-regulated genes in wild-type peritoneal macrophages versus total genes on the array (20k genes). The position of the miR-155 native seed sequence is indicated as "N", and the 3 possible edited seeds as E1, E2 and E3. The blue shaded area represents the bottom 90% of the 4,096 hexamer RNA motifs based on the rank position of P values. (B) The same analysis was performed as in (A) but in bic/miR-155 deficient peritoneal macrophages. (C) Seed enrichment analysis in wild-type fibroblasts. (D) Seed enrichment analysis in bic/miR-155 deficient fibroblasts.

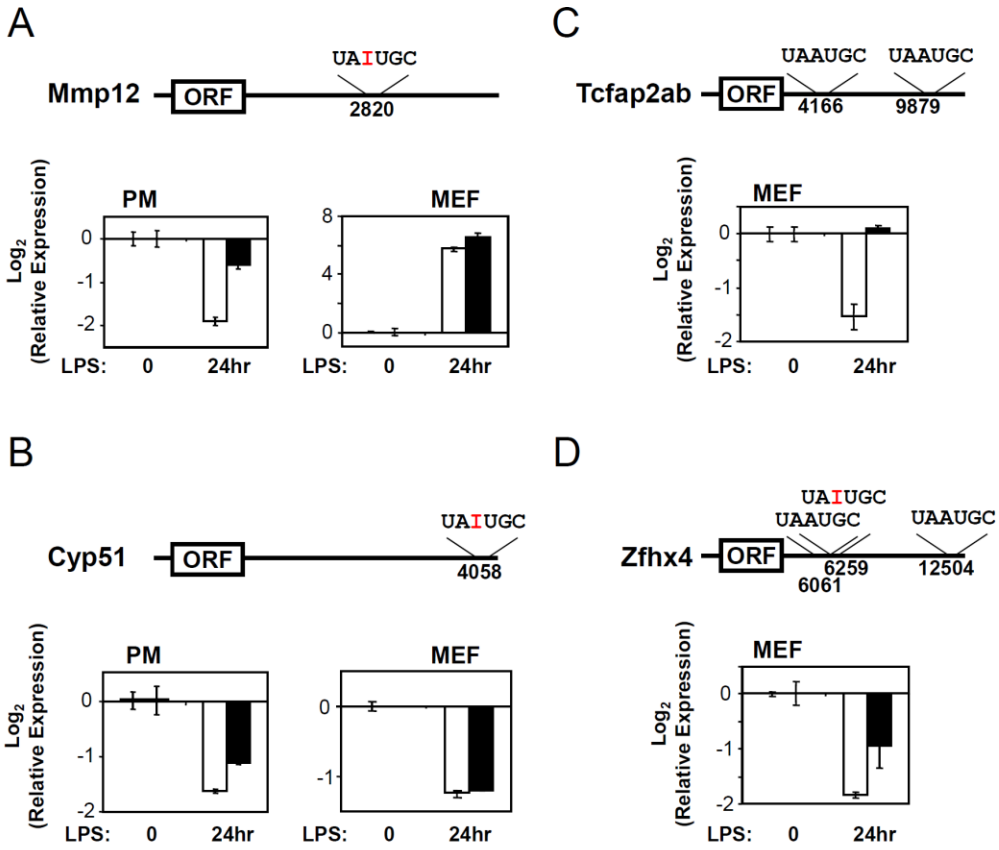


Figure 4.4 mRNA target of edited miR-155 has reduced gene repression in bic/miR-155-deficient cells

(A) Relative mRNA expression of an edited miR-155 target gene Mmp12 in wt (white) and bic/miR-155-deficient (black) PMs and MEFs were analyzed by RT-qPCR. The relative position of an edited seed sequence miR-155_E1 (UAIUGC) within the 3'UTR is indicated. (B) Relative mRNA expression of Cyp51 in wt (white) and bic/miR-155-deficient (black) PMs and MEFs were analyzed by RT-qPCR. A miR-155_E1 seed sequence is indicated within the 3'-UTR. (C)(D) Relative mRNA expression of native miR-155 target genes Tcfap2ab and Zfhx4 in wt (white) and bic/miR-155-deficient (black) MEFs were analyzed by RT-qPCR. Native miR-155 (UAAUGC) and edited miR-155_E1 seed sequences are indicated within the 3'UTR.

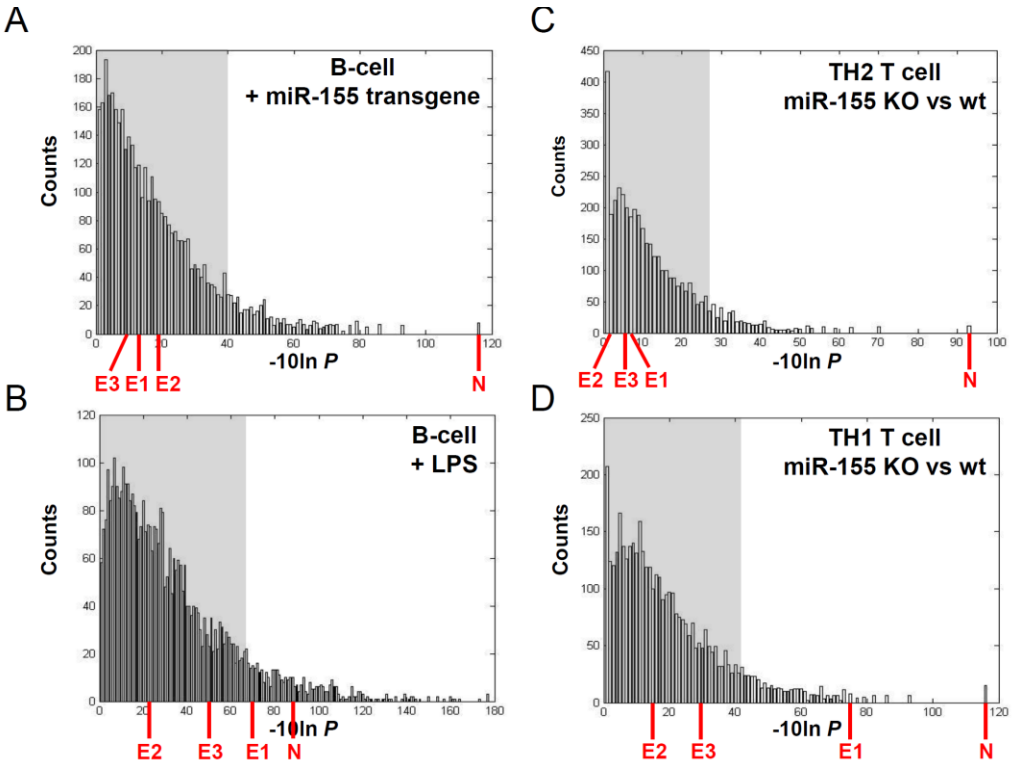


Figure 4.5 Gene repression by edited miR-155 in activated pro-inflammatory adaptive immune cells

(A) Seed enrichment analysis was applied to down-regulated genes in B-cells overexpressing a miR-155 transgene versus genes without expression change using microarray data from Costinean et al. (Costinean et al., 2006). The position of native and edited miR-155 seed sequences is indicated as in Figure 3. (B) Seed enrichment analysis was performed with LPS down-regulated genes in B-cells versus genes without expression change (Zhu et al., 2004). (C) Seed enrichment analysis was applied to hyper-expressed genes versus genes without expression change in bic/miR-155 deficient TH2 T cells as compared to wild-type TH2 cells (Rodriguez et al., 2007). (D) In hyper-expressed genes in bic/miR-155 deficient TH1 T cells (Rodriguez et al., 2007) were used for the seed enrichment analysis.

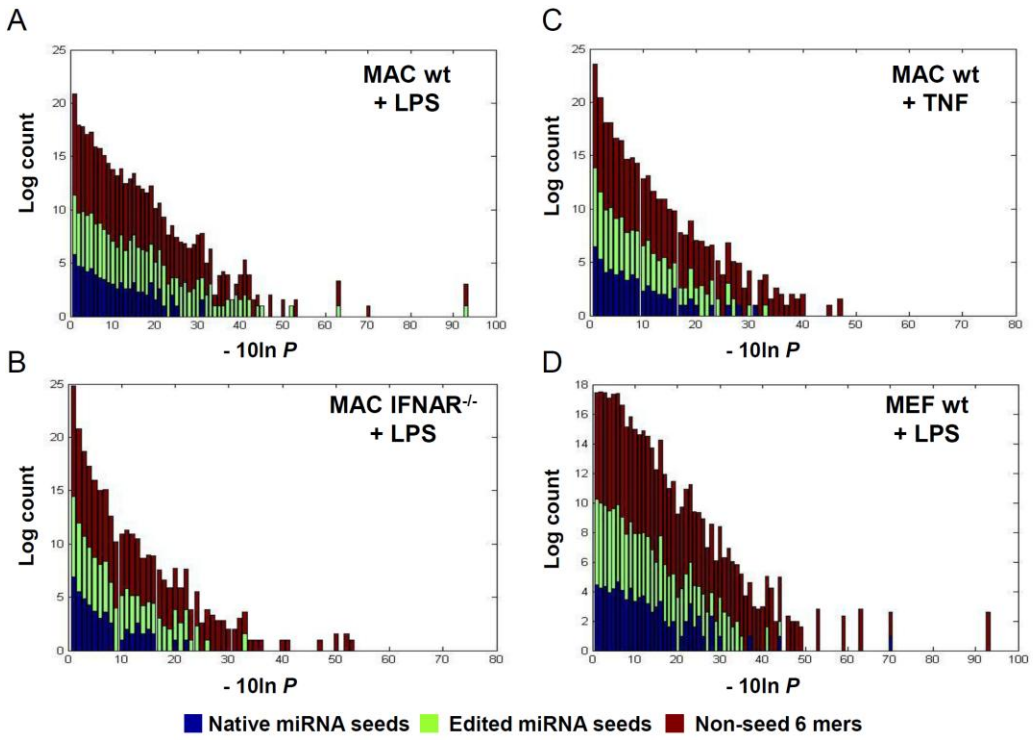


Figure 4.6 Gene repression by widespread miRNA editing is Type I IFN-dependent (A) Seed enrichment analysis was performed with LPS down-regulated genes in wild-type peritoneal macrophages. All 4,096 possible hexamer RNA motifs were classified into non-edited native miRNA seeds (shown in blue), potentially edited miRNA seeds (shown in green), and non miRNA seed hexamers that do not conform to either (shown in brick red). (B) Seed enrichment analysis was applied to LPS down-regulated genes in IFNAR$^{-/-}$ peritoneal macrophages. (C) Seed enrichment analysis was performed with TNF down-regulated genes in wild-type peritoneal macrophages. (D) Seed enrichment analysis was performed with LPS down-regulated genes in wild-type MEFs.

A

miRNAs induced by inflammatory stimuli	Reference	Derived edited seed	Edited Seed ID	Percentile
miR-351	Pedersen	CCCUGI	miR-351_E1	< 0.5%
miR-1	Pedersen	GGIIUG	miR-1_E1	< 0.5%
miR-30a-3p, miR-30e*	Pedersen	UUUCIG	miR-30a-3p_E1, miR-30e*	1.5%
miR-132	Taganov	IICIGU	miR-132_E7	3.0%
miR-448	Pedersen	UGCIUI	miR-448_E3	3.4%
miR-146	Taganov	GIGAIC	miR-146_E2	3.8%
miR-30a-5p, miR-30e, miR-30d, miR-30c, miR-30b	Pedersen	GUIIAC	miR-30a-5p_E6, miR-30b,c,d,e_E6	6.4%
miR-128a	Pedersen	CACIGU	miR-128_E2	7.0%
miR-155	O'Connell	UAIUGC	miR-155_E1	7.2%

B

Constitutively expressed miRNAs	Reference	Derived edited seed	Edited Seed ID	Percentile
miR-125a,b	Liu	CCCUGI	miR-125a,b_E1	< 0.5%
let-7	Liu	GIGGUI	let-7_E2	< 0.5%
miR-130	Liu	AGUGCI	miR-130_E1	< 0.5%
miR-7	Liu	GGIIGI	miR-7_E4	0.7%
miR-214	Liu	CIGCIG	miR-214_E1	0.8%
miR-20	Liu	AAIGUG	miR-20_E1	3.6%

Figure 4.7 miRNAs that has edited seed been enriched

(A) Table of miRNAs previously shown to be induced in response to a number of inflammatory stimuli and that are related to derived edited seed sequences that were enriched (within the top 10% rank position among all 4,096 hexamers) in the 3'UTR of LPS down-regulated genes in wild-type peritoneal macrophages in (Figure 4.5A). (B) Table of miRNAs that are constitutively expressed in RAW cells and related to derived edited seed sequences enriched in wild-type peritoneal macrophages in (Figure 4.5A).

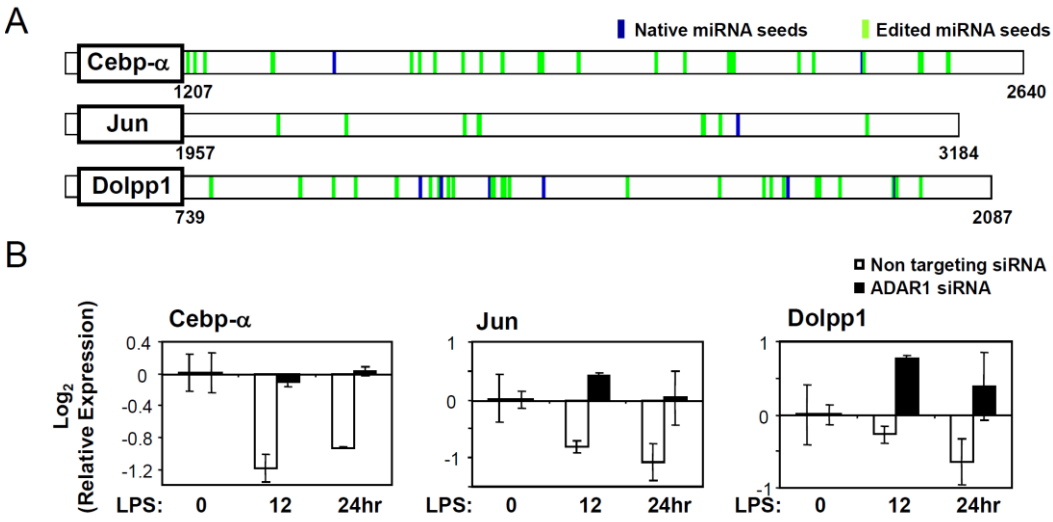
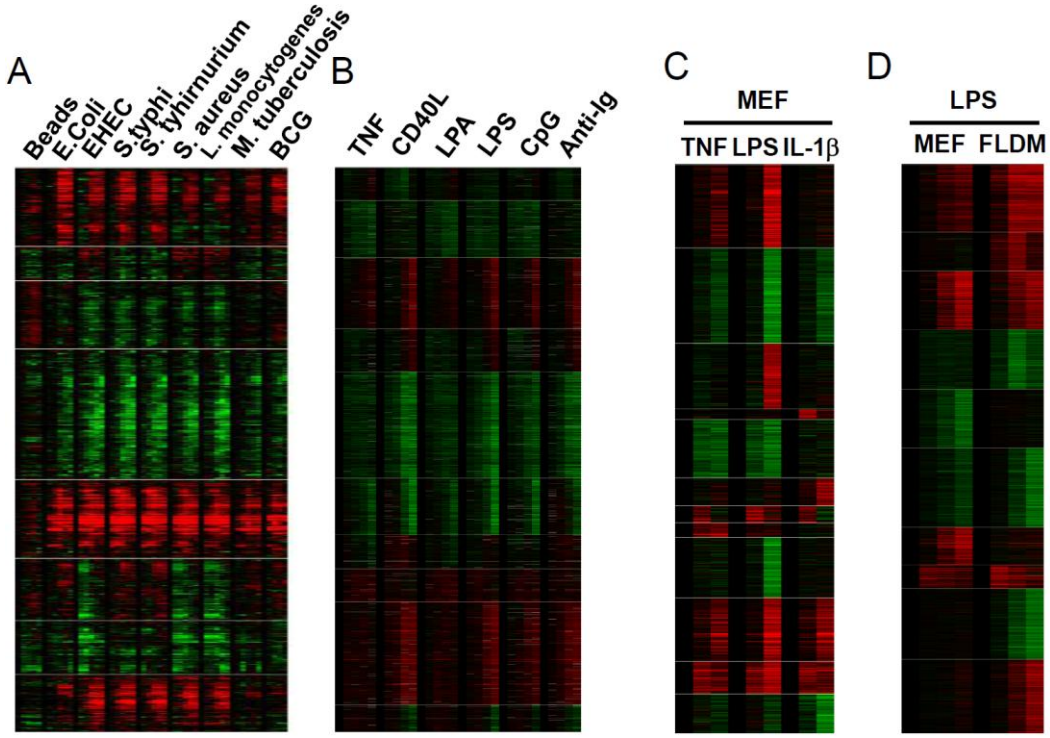
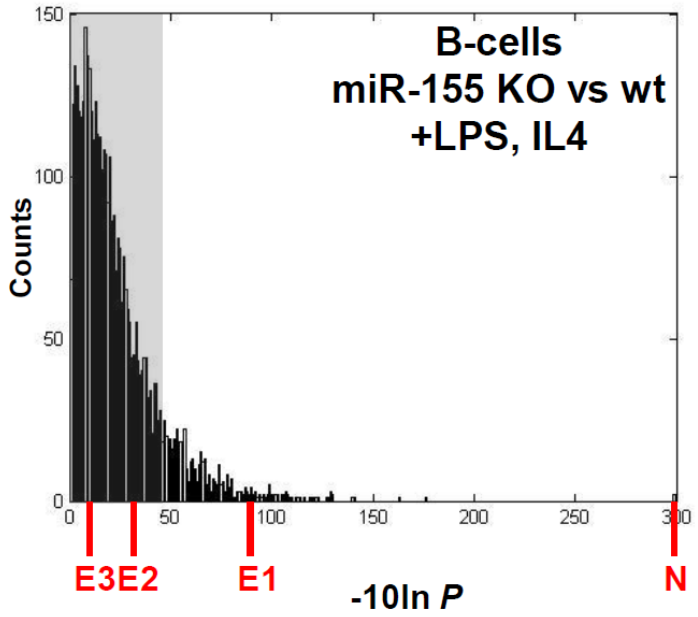


Figure 4.8 miRNA editing is ADAR-1L dependent

(A) Schematic of the mRNAs of three genes, indicating potential target sequences for native (blue) and edited (green) miRNA seed sequences. Seed sequences that are statistically enriched in LPS-induced gene repression programs in wild type but not IFNAR$^{-/-}$ macrophages are indicated. Native and edited seed sequences shown in the 3'-UTRs are only those derived from miRNAs validated by traditional cloning methods and/or northern blot.. (B) Relative mRNA expression of Cebp-α, Jun



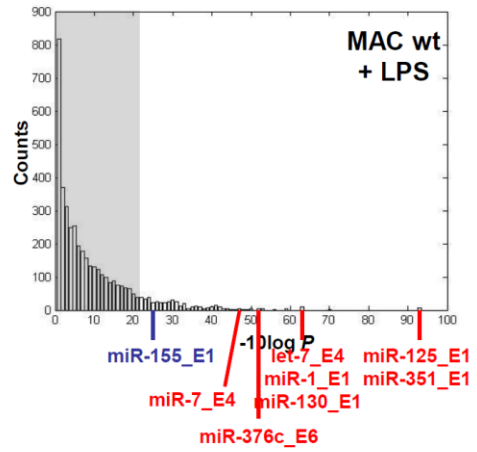
Appendix Figure 4.1 Inflammatory responses involve gene repression programs
(A) Gene activation and gene repression programs elicited in human macrophages in response to latex beads, or infection with E. Coli, EHEC, S. typhi, S. tyhirnurium, S. aureus, L. monocytogenes, M. tuberculosis and BCG at 0, 1, 2, 6, 12 and 24 hours, and revealed by K-means clustering of data produced by Nau et al. (Nau et al., 2002). (B) K-means clustering of gene expression programs in splenic B-cells stimulated with TNF, CD40L, LPA, LPS, CpG and Anti-Ig at 0, 0.5, 1, 2, 4 hours using microarray data from Zhu et al. (Zhu et al., 2004) (C) Stimulus-specific gene activation and gene repression programs in MEFs in response to TNF (1 ng/ml), LPS (0.1 μg/ml) and IL-1β (1 ng/ml) at 0, 1, 8 hours as revealed by K-means clustering. (D) Cell-type specific gene activation and gene repression programs in MEFs and FLDMs stimulated with LPS (0.1μg/ml) for 0, 1, 3, 8 hrs as revealed by K-means clustering.



Appendix Figure 4.2 Gene repression by edited miR-155 in activated B-cells

Seed enrichment analysis was applied to hyper-expressed genes versus no-change genes in bic/miR-155 deficient splenic B-cells as compared to wild-type cells. (Microarray data were from Vigorito et al.(Rodriguez et al., 2007)) Cells were stimulated with LPS and IL-4 for 24 hours. The position of the miR-155 native seed sequence is indicated as "N", the 3 possible edited seeds as E1, E2 and E3. The blue shaded area represents 90% of the 4,096 hexamer RNA motifs.

A

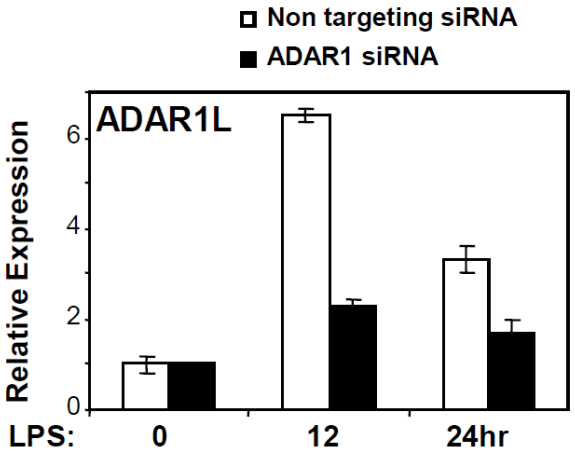


B

Edited microRNA seed ID	Seed	-10ln P
miR-351_E1	CCCUGI	92.1
miR-1_E1	GGIIUG	62.1
miR-30a-3p_E1, miR-30e*	UUUCIG	40.7
miR-132_E7	IICIGU	33.2
miR-448_E3	UGCIUI	32.1
miR-146_E2	GIGAIC	31.0
miR-30a-5p_E6, miR-30b,c,d,e_E6	GUIIAC	26.6
miR-128_E2	CACIGU	25.9
miR-155_E1	CCCUGI	25.5
miR-125a,b_E1	CCCUGI	92.1
let-7_E2	GIGGUI	62.1
miR-130_E1	AGUGCI	62.1
miR-7_E4	GGIIGI	49.6
miR-214_E1	CIGCIG	47.1
miR-20_E1	AAIGUG	31.7

Appendix Figure 4.3 Edited seeds related to inflammatory stimuli induced miRNAs and constitutively expressed miRNAs

(A) Seed enrichment analysis was performed with LPS-repressed genes in wild-type peritoneal macrophages as shown in Figure 5a. Bars with specific edited seed sequence IDs indicate the rank positions of some of the edited seeds deriving from miRNAs that are either highly expressed or induced by inflammatory stimuli. (B) Table of edited microRNA seed sequences with -10ln P values corresponding to Fig. 5e and 5f.



Appendix Figure 4.4 siRNA knockdown of ADAR1 in peritoneal macrophages
Relative mRNA expression of ADAR-1L in PMs with ADAR1 and non-targeting siRNA knockdown were measured with RT-qPCR. Cells were stimulated with LPS (0.1 µg/ml) for 12 and 24 hrs.

Appendix Table 4.1Rank sum test results for miR155-derived native and edited seed sequences in various immune cells

		Fig 3a	Fig 3b	Fig 3c	Fig 3d	Fig 4a	Fig 4b	Fig 4c	Fig 4d	Suppl. Fig 2
Name	Seed	MAC wt +LPS	MAC miR-155 KO +LPS	MEF wt +LPS	MEF miR-155 KO + LPS	B-cell + miR-155 transgene	B-cell + LPS	TH1 T cell miR-155 KO vs wt	TH2 T cell miR-155 KO vs wt	B-cells miR-155 KO vs wt + LPS, IL4
miR155 N	GCATTA	17.0%	36.4%	4%	21%	<0.5%	5%	<0.5%	<0.5%	5%
miR155 E1	GCACTA	4%	20%	37%	47%	51%	9%	1%	59%	8%
miR155 E2	GCATCA	62%	40%	69%	46%	64%	20%	22%	65%	63%
miR155 E3	GCACCA	56%	28%	78%	91%	36%	54%	50%	83%	24%

Appendix Table 4.2 Rank sum test results for miR155-derived native and edited seed sequences

Name	Sequence	B-cell + miR-155 transgene	B-cell + LPS	TH1 T cell miR-155 KO vs wt	TH2 T cell miR-155 KO vs wt	B-cells miR-155 KO vs wt + LPS, IL4
miR155 N/ 6mer (p1-6)	CATTAA	3%	35%	<0.5%	<0.5%	5%
miR155 E1/ 6mer (p1-6)	CACTAA	4%	11%	7%	70%	5%
miR155 E2/ 6mer (p1-6)	CATCAA	49%	28%	55%	88%	74%
miR155 E3/ 6mer (p1-6)	CACCAA	82%	64%	43%	89%	92%
miR-155 N/ 7mer (p1-7)	GCATTAA	<0.5%	31%	<0.5%	<0.5%	<0.5%
miR-155 E1/ 7mer (p1-7)	GCACTAA	41%	8%	<0.5%	51%	11%
miR-155 E2/ 7mer (p1-7)	GCATCAA	68%	6%	98%	94%	56%
miR-155 E3/ 7mer (p1-7)	GCACCAA	31%	66%	26%	96%	75%
miR155 N/ A + 6mer (p2-7)	GCATTAA	<0.5%	33%	5%	<0.5%	<0.5%
miR155 E1/ A + 6mer (p2-7)	GCACTAA	47%	9%	5%	52%	10%
miR155 E2/ A + 6mer (p2-7)	GCATCAA	73%	7%	98%	93%	56%
miR155 E3/ A + 6mer (p2-7)	GCACCAA	37%	69%	28%	95%	75%
miR-155 N/ 7mer (p2-8)	AGCATTA	<0.5%	1%	<0.5%	<0.5%	<0.5%
miR-155 E1/ 7mer (p2-8)	AGCACTA	95%	70%	7%	30%	14%
miR-155 E2/ 7mer (p2-8)	AGCATCA	72%	32%	5%	37%	66%
miR-155 E3/ 7mer (p2-8)	AGCACCA	90%	37%	24%	99%	51%

Appendix Table 4.3 Potential edited miRNA target genes

Accession	Gene	Edited Seed Count (Total Sanger9.1)	Edited Seed Count (Cloned/Northern validated)
NM_007678	CCAAT/enhancer binding protein (C/EBP), alpha	69	36
NM_007711	chloride channel 3	31	16
NM_007757	coproporphyrinogen oxidase	24	17
NM_007760	carnitine acetyltransferase	81	39
NM_007891	E2F transcription factor 1	27	14
NM_007918	eukaryotic translation initiation factor 4E binding protein 1	29	20
NM_007960	ets variant gene 1	16	11
NM_008131	glutamate-ammonia ligase (glutamine synthase)	34	17
NM_008250	H2.0-like homeo box gene	11	7
NM_008488	Rho guanine nucleotide exchange factor (GEF) 1	25	13
NM_008787	pericentrin 2	15	8
NM_008863	protein kinase inhibitor beta, cAMP dependent, testis specific	26	14
NM_008992	ATP-binding cassette, sub-family D (ALD), member 4	14	8
NM_009053	radical fringe gene homolog (Drosophila)	60	40
NM_009447	tubulin, alpha 4	19	11
NM_009517	wild-type p53-induced gene 1	157	95
NM_009624	adenylate cyclase 9	10	4
NM_009741	B-cell leukemia/lymphoma 2	111	66
NM_009870	cyclin-dependent kinase 4	10	7
NM_009924	cannabinoid receptor 2 (macrophage)	51	33
NM_010121	eukaryotic translation initiation factor 2 alpha kinase 3	24	13
NM_010150	nuclear receptor subfamily 2, group F, member 6	18	11
NM_010286	glucocorticoid-induced leucine zipper	35	18
NM_010347	amino-terminal enhancer of split	30	22
NM_010365	general transcription factor II I	41	27
NM_010566	inositol polyphosphate-5-phosphatase D	47	29
NM_010591	Jun oncogene	23	10
NM_010598	shaker-related subfamily, beta member 2	98	59
NM_010631	kinesin family member C3	32	19
NM_010638	basic transcription element binding protein 1	47	24
NM_010735	lymphotoxin A	23	16
NM_010784	midkine	14	8
NM_010957	8-oxoguanine DNA-glycosylase 1	14	10
NM_011250	retinoblastoma-like 2	21	12
NM_011424	nuclear receptor co-repressor 2	51	25
NM_011467	sepiapterin reductase	17	14
NM_011549	transcription factor EB	37	22

Appendix Table 4.4 Potential edited miRNA target genes continued

NM_011697	vascular endothelial growth factor B	21	12
NM_011841	mitogen-activated protein kinase 7	15	8
NM_011906	transmembrane domain protein regulated in adipocytes	34	17
NM_011937	glucosamine-6-phosphate deaminase 1	39	25
NM_013849	dual specificity phosphatase 13	19	6
NM_015763	lipin 1	59	38
NM_015818	heparan sulfate 6-O-sulfotransferase 1	84	56
NM_016713	mitogen-activated protein kinase kinase kinase kinase 6	52	34
NM_016845	proacrosin binding protein	13	8
NM_016920	ATPase, H+ transporting, lysosomal V0 subunit a isoform 1	65	34
NM_017380	septin 9	86	53
NM_020329	Dolpp1 dolichyl pyrophosphate phosphatase 1	57	33
NM_020517	lens epithelial protein	48	27
NM_021528	carbohydrate sulfotransferase 12	21	14
NM_021555	brain protein 16	10	7
NM_021567	poly(rC) binding protein 4	29	20
NM_023166	mammary tumor virus receptor 2	18	12
NM_023792	pantothenate kinase 1	25	11
NM_025320	RIKEN cDNA 0610012G03 gene	14	7
NM_025479	RIKEN cDNA 2810021B07 gene	18	6
NM_026189	Eepd1 endonuclease/exonuclease/phosphatase family domain containing 1	10	5
NM_026302	dynactin 4	33	16
NM_026371	Loh12cr1 loss of heterozygosity, 12, chromosomal region 1 homolog	39	18
NM_026423	RIKEN cDNA 2410018C20 gene	11	7
NM_026514	CDC42 effector protein (Rho GTPase binding) 3	11	7
NM_026524	Mid1ip1 Mid1 interacting protein 1 (gastrulation specific G12-like (zebrafish))	32	17
NM_026633	RIKEN cDNA 9530058B02 gene	14	10
NM_028133	EGL nine homolog 3 (C. elegans)	25	17
NM_028732	RIKEN cDNA 4632428N05 gene	111	66
NM_029091	Klc4 kinesin light chain 4	11	8
NM_030258	cDNA sequence BC003323	106	61
NM_030714	deltex 3 homolog (Drosophila)	38	26
NM_031251	cystinosis, nephropathic	48	24
NM_033072	DNA segment, Chr 10, Wayne State University 93, expressed	32	19
NM_033134	inositol polyphosphate-5-phosphatase E	38	23

Chapter 5 : Conclusions

CONCLUSIONS

Through the combination of system-wide "top-down" and mechanistic "bottom-up" approaches, we classified the inflammatory expression response into seven (?) separable gene programs that are functional targets of NFκB, bZIP and IRF/ISGF3, alone or in combination. The surprisingly simple classification of genes seems to largely explain LPS-responsive inflammatory gene expression observed within the first eight hours of a timecourse. We did not observe a class of genes that are regulated by synergistic activation of two different transcription factors, instead, we discovered a new class of gene expression that are controlled by synergistic AND gate between regulated mRNA decay and transcription. Our results demonstrated an example of how stimulus-induced regulation of mRNA stability can play an important role in determining stimulus-specific expression and pathogen specific responses.

We have demonstrated that three major stimulus-inducible transcription factors, NFκB, bZIP and IRF/ISGF3 are responsible for activating the LPS-responsive transcriptional program. However, we did not examine the functional connection of these stimulus-inducible transcriptional regulators in the context of constitutively expressed, cell type-specific master regulators (eg. PU.1 in macrophages) that define the cell type-specific transcriptional landscape (Ghisletti et al.; Heinz et al.). Cell type specific master regulators prepare the chromatin from a repressive state to a poised state that can be immediately activated upon stimulation (Natoli). We also did not consider the potential

chromatin changes that might happen at later times post stimulation (24 hours or 48 hours). These are important questions that can be addressed by examining the effect of inducible transcription factors (in knock out cells) on chromatin states in unstimulated cells as well as late time points.

We identify that the NFκB p50 homodimer cross-regulates the IRF and anti-viral response by directly binding to the G-IRE sequences at the promoter of IFN responsive genes. We utilized microarray analyses, motif search and statistical analyses to identify the specific subset of IREs that contain guanine-rich sequences flanking the IRE motif. Mathematical modeling allowed us to characterize the potential functional consequences of p50 homodimer repressing a composite promoter that is a synergistic AND gate between NFκB and IRF. Our model predicted that the p50 homodimer may enforce stimulus-specific repression of composite promoters. Furthermore, we show the expression of a composite promoter, the anti-viral regulator IFNβ was found to be stimulus-restricted by p50 homodimer binding to the G-IRE-containing enhancer to suppress cytotoxic IFN signaling that might be harmful to the cells. Our results demonstrated the novel role of NFkB p50 homodimer in cross-regulating IREs and the anti-viral response.

The innate immune response induces not only a gene activation program but also a gene repression program of hundreds of genes, which are often times been ignored. By characterizing the possibility of gene down-regulation by miRNAs, we employed a computational strategy that examined not only native miRNA seeds but also potential A-

to-I edited miRNA seed sequences. To our surprise, we identified a specific A-to-I edited seed sequence of the NFκB inducible miR-155 to be highly enriched in the 3'UTR of gene repression programs of activated macrophages, B-cells and inflammatory TH1 cells, but not in anti-inflammatory TH2 cells or non-immune cells. The enrichment of the edited seed sequence also correlates with the induction of the IFN pathway and the expression of the deaminase, ADAR1. Many of the edited seeds of constitutive or stimulus induced miRNAs are also enriched in the repression program in LPS induced macrophages. Our work suggests that when characterizing functional targets of miRNAs, it is important to consider cell type and stimulus specific post-transcriptional miRNA modifications. A major challenge to extend these studies is to develop methods to reliably discover small RNA modification by refined cloning and sequencing methods or by mass spectrometry.

REFERENCES

Akira, S., and Takeda, K. (2004). Toll-like receptor signalling. Nat Rev Immunol *4*, 499-511.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen recognition and innate immunity. Cell *124*, 783-801.

Al-Souhibani, N., Al-Ahmadi, W., Hesketh, J.E., Blackshear, P.J., and Khabar, K.S. The RNA-binding zinc-finger protein tristetraprolin regulates AU-rich mRNAs involved in breast cancer-related processes. Oncogene *29*, 4205-4215.

Al-Souhibani, N., Al-Ahmadi, W., Hesketh, J.E., Blackshear, P.J., and Khabar, K.S. (2010). The RNA-binding zinc-finger protein tristetraprolin regulates AU-rich mRNAs involved in breast cancer-related processes. Oncogene *29*, 4205-4215.

Amit, I., Garber, M., Chevrier, N., Leite, A.P., Donner, Y., Eisenhaure, T., Guttman, M., Grenier, J.K., Li, W., Zuk, O., *et al.* (2009). Unbiased reconstruction of a mammalian transcriptional network mediating pathogen responses. Science *326*, 257-263.

Badis, G., Berger, M.F., Philippakis, A.A., Talukder, S., Gehrke, A.R., Jaeger, S.A., Chan, E.T., Metzler, G., Vedenko, A., Chen, X., *et al.* (2009). Diversity and complexity in DNA recognition by transcription factors. Science *324*, 1720-1723.

Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., and Pasquinelli, A.E. (2005). Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. Cell *122*, 553-563.

Basak, S., Kim, H., Kearns, J.D., Tergaonkar, V., O'Dea, E., Werner, S.L., Benedict, C.A., Ware, C.F., Ghosh, G., Verma, I.M., *et al.* (2007). A fourth IkappaB protein within the NF-kappaB signaling module. Cell *128*, 369-381.

Beinke, S., Deka, J., Lang, V., Belich, M.P., Walker, P.A., Howell, S., Smerdon, S.J., Gamblin, S.J., and Ley, S.C. (2003). NF-kappaB1 p105 negatively regulates TPL-2 MEK kinase activity. Mol Cell Biol *23*, 4739-4752.

Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., Kuhlman, T., and Phillips, R. (2005a). Transcriptional regulation by the numbers: applications. Curr Opin Genet Dev *15*, 125-135.

Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. (2005b). Transcriptional regulation by the numbers: models. Curr Opin Genet Dev *15*, 116-124.

Blow, M.J., Grocock, R.J., van Dongen, S., Enright, A.J., Dicks, E., Futreal, P.A., Wooster, R., and Stratton, M.R. (2006). RNA editing of human microRNAs. Genome Biol *7*, R27.

Bohuslav, J., Kravchenko, V.V., Parry, G.C., Erlich, J.H., Gerondakis, S., Mackman, N., and Ulevitch, R.J. (1998). Regulation of an essential innate immune response by the p50 subunit of NF-kappaB. J Clin Invest *102*, 1645-1652.

Borden, E.C., Sen, G.C., Uze, G., Silverman, R.H., Ransohoff, R.M., Foster, G.R., and Stark, G.R. (2007). Interferons at age 50: past, current and future impact on biomedicine. Nat Rev Drug Discov *6*, 975-990.

Bradley, J.R. (2008). TNF-mediated inflammatory disease. J Pathol *214*, 149-160.

Braganca, J., Genin, P., Bandu, M.T., Darracq, N., Vignal, M., Casse, C., Doly, J., and Civas, A. (1997). Synergism between multiple virus-induced factor-binding elements involved in the differential expression of interferon A genes. J Biol Chem *272*, 22154-22162.

Brivanlou, A.H., and Darnell, J.E., Jr. (2002). Signal transduction and the control of gene expression. Science *295*, 813-818.

Brook, M., Tchen, C.R., Santalucia, T., McIlrath, J., Arthur, J.S., Saklatvala, J., and Clark, A.R. (2006). Posttranslational regulation of tristetraprolin subcellular localization and protein stability by p38 mitogen-activated protein kinase and extracellular signal-regulated kinase pathways. Mol Cell Biol *26*, 2408-2418.

Buchler, N.E., Gerland, U., and Hwa, T. (2003). On schemes of combinatorial transcription logic. Proc Natl Acad Sci U S A *100*, 5136-5141.

Calin, G.A., and Croce, C.M. (2006). MicroRNA signatures in human cancers. Nat Rev Cancer *6*, 857-866.

Carey, M., Lin, Y.S., Green, M.R., and Ptashne, M. (1990). A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. Nature *345*, 361-364.

Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., Frisch, M., Bayerlein, M., and Werner, T. (2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics *21*, 2933-2942.

Caudy, A.A., Ketting, R.F., Hammond, S.M., Denli, A.M., Bathoorn, A.M., Tops, B.B., Silva, J.M., Myers, M.M., Hannon, G.J., and Plasterk, R.H. (2003). A micrococcal nuclease homologue in RNAi effector complexes. Nature *425*, 411-414.

Chang, M., Lee, A.J., Fitzpatrick, L., Zhang, M., and Sun, S.C. (2009). NF-kappa B1 p105 regulates T cell homeostasis and prevents chronic inflammation. J Immunol *182*, 3131-3138.

Chen, F.E., Huang, D.B., Chen, Y.Q., and Ghosh, G. (1998a). Crystal structure of p50/p65 heterodimer of transcription factor NF-kappaB bound to DNA. Nature *391*, 410-413.

Chen, Y.Q., Ghosh, S., and Ghosh, G. (1998b). A novel DNA recognition mode by the NF-kappa B p65 homodimer. Nat Struct Biol *5*, 67-73.

Chow, E.K., O'Connell R, M., Schilling, S., Wang, X.F., Fu, X.Y., and Cheng, G. (2005). TLR agonists regulate PDGF-B production and cell proliferation through TGF-beta/type I IFN crosstalk. EMBO J *24*, 4071-4081.

Cogswell, P.C., Scheinman, R.I., and Baldwin, A.S., Jr. (1993). Promoter of the human NF-kappa B p50/p105 gene. Regulation by NF-kappa B subunits and by c-REL. J Immunol *150*, 2794-2804.

Costinean, S., Zanesi, N., Pekarsky, Y., Tili, E., Volinia, S., Heerema, N., and Croce, C.M. (2006). Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. Proc Natl Acad Sci U S A *103*, 7024-7029.

Dahlberg, J.E., and Lund, E. (2007). Micromanagement during the innate immune response. Sci STKE *2007*, pe25.

Decker, T., Muller, M., and Stockinger, S. (2005). The yin and yang of type I interferon activity in bacterial infection. Nat Rev Immunol *5*, 675-687.

Doench, J.G., and Sharp, P.A. (2004). Specificity of microRNA target selection in translational repression. Genes Dev *18*, 504-511.

Doyle, S., Vaidya, S., O'Connell, R., Dadgostar, H., Dempsey, P., Wu, T., Rao, G., Sun, R., Haberland, M., Modlin, R., *et al.* (2002). IRF3 mediates a TLR3/TLR4-specific antiviral gene program. Immunity *17*, 251-263.

Eichler, G.S., Huang, S., and Ingber, D.E. (2003). Gene Expression Dynamics Inspector (GEDI): for integrative analysis of expression profiles. Bioinformatics *19*, 2321-2322.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A *95*, 14863-14868.

Emmons, J., Townley-Tilson, W.H., Deleault, K.M., Skinner, S.J., Gross, R.H., Whitfield, M.L., and Brooks, S.A. (2008). Identification of TTP mRNA targets in human

dendritic cells reveals TTP as a critical regulator of dendritic cell maturation. RNA *14*, 888-902.

Escalante, C.R., Nistal-Villan, E., Shen, L., Garcia-Sastre, A., and Aggarwal, A.K. (2007). Structure of IRF-3 bound to the PRDIII-I regulatory element of the human interferon-beta enhancer. Mol Cell *26*, 703-716.

Fan, C.M., and Maniatis, T. (1989). Two different virus-inducible elements are required for human beta-interferon gene regulation. EMBO J *8*, 101-110.

Farh, K.K., Grimson, A., Jan, C., Lewis, B.P., Johnston, W.K., Lim, L.P., Burge, C.B., and Bartel, D.P. (2005). The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. Science *310*, 1817-1821.

Foster, S.L., Hargreaves, D.C., and Medzhitov, R. (2007). Gene-specific control of inflammation by TLR-induced chromatin modifications. Nature *447*, 972-978.

Fujii, Y., Shimizu, T., Kusumoto, M., Kyogoku, Y., Taniguchi, T., and Hakoshima, T. (1999). Crystal structure of an IRF-DNA complex reveals novel DNA recognition and cooperative binding to a tandem repeat of core sequences. EMBO J *18*, 5028-5041.

Garrick, R.A., Polefka, T.G., Cua, W.O., and Chinard, F.P. (1986). Water permeability of alveolar macrophages. Am J Physiol *251*, C524-528.

George, C.X., and Samuel, C.E. (1999). Human RNA-specific adenosine deaminase ADAR1 transcripts possess alternative exon 1 structures that initiate from different promoters, one constitutively active and the other interferon inducible. Proc Natl Acad Sci U S A *96*, 4621-4626.

Ghisletti, S., Barozzi, I., Mietton, F., Polletti, S., De Santa, F., Venturini, E., Gregory, L., Lonie, L., Chew, A., Wei, C.L., *et al.* Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. Immunity *32*, 317-328.

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

Gilchrist, M., Thorsson, V., Li, B., Rust, A.G., Korb, M., Roach, J.C., Kennedy, K., Hai, T., Bolouri, H., and Aderem, A. (2006). Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. Nature *441*, 173-178.

Giorgetti, L., Siggers, T., Tiana, G., Caprara, G., Notarbartolo, S., Corona, T., Pasparakis, M., Milani, P., Bulyk, M.L., and Natoli, G. Noncooperative interactions between

transcription factors and clustered DNA binding sites enable graded transcriptional responses to environmental inputs. Mol Cell *37*, 418-428.

Goriely, S., Molle, C., Nguyen, M., Albarani, V., Haddou, N.O., Lin, R., De Wit, D., Flamand, V., Willems, F., and Goldman, M. (2006). Interferon regulatory factor 3 is involved in Toll-like receptor 4 (TLR4)- and TLR3-induced IL-12p35 gene activation. Blood *107*, 1078-1084.

Habig, J.W., Dale, T., and Bass, B.L. (2007). miRNA editing--we should have inosine this coming. Mol Cell *25*, 792-793.

Hao, S., and Baltimore, D. (2009). The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules. Nat Immunol *10*, 281-288.

Hartner, J.C., Schmittwolf, C., Kispert, A., Muller, A.M., Higuchi, M., and Seeburg, P.H. (2004). Liver disintegration in the mouse embryo caused by deficiency in the RNA-editing enzyme ADAR1. J Biol Chem *279*, 4894-4902.

Hawkins, R.D., Hon, G.C., and Ren, B. Next-generation genomics: an integrative approach. Nat Rev Genet *11*, 476-486.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. 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., Leung, T.H., and Baltimore, D. (2003). Genetic analysis of NF-kappaB/Rel transcription factors defines functional specificities. EMBO J *22*, 5530-5539.

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

Hsiao, A., Ideker, T., Olefsky, J.M., and Subramaniam, S. (2005). VAMPIRE microarray suite: a web-based platform for the interpretation of gene expression data. Nucleic Acids Res *33*, W627-632.

Iyer, V.R., Eisen, M.B., Ross, D.T., Schuler, G., Moore, T., Lee, J.C., Trent, J.M., Staudt, L.M., Hudson, J., Jr., Boguski, M.S., *et al.* (1999). The transcriptional program in the response of human fibroblasts to serum. Science *283*, 83-87.

Jackson, R.J., and Standart, N. (2007). How do microRNAs regulate gene expression? Sci STKE *2007*, re1.

John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human MicroRNA targets. PLoS Biol *2*, e363.

Kang, S.M., Tran, A.C., Grilli, M., and Lenardo, M.J. (1992). NF-kappa B subunit regulation in nontransformed CD4+ T lymphocytes. Science *256*, 1452-1456.

Karin, M., Lawrence, T., and Nizet, V. (2006). Innate immunity gone awry: linking microbial infections to chronic inflammation and cancer. Cell *124*, 823-835.

Kawahara, Y., Zinshteyn, B., Chendrimada, T.P., Shiekhattar, R., and Nishikura, K. (2007a). RNA editing of the microRNA-151 precursor blocks cleavage by the Dicer-TRBP complex. EMBO Rep *8*, 763-769.

Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A.G., and Nishikura, K. (2007b). Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. Science *315*, 1137-1140.

Kawai, T., and Akira, S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. Nat Immunol *11*, 373-384.

Kim, H.D., Shay, T., O'Shea, E.K., and Regev, A. (2009). Transcriptional regulatory circuits: predicting numbers from alphabets. Science *325*, 429-432.

Koenig Merediz, S.A., Schmidt, M., Hoppe, G.J., Alfken, J., Meraro, D., Levi, B.Z., Neubauer, A., and Wittig, B. (2000). Cloning of an interferon regulatory factor 2 isoform with different regulatory ability. Nucleic Acids Res *28*, 4219-4224.

Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., *et al.* (2005). Combinatorial microRNA target predictions. Nat Genet *37*, 495-500.

Krutzfeldt, J., Rajewsky, N., Braich, R., Rajeev, K.G., Tuschl, T., Manoharan, M., and Stoffel, M. (2005). Silencing of microRNAs in vivo with 'antagomirs'. Nature *438*, 685-689.

Kunsch, C., Ruben, S.M., and Rosen, C.A. (1992). Selection of optimal kappa B/Rel DNA-binding motifs: interaction of both subunits of NF-kappa B with DNA is required for transcriptional activation. Mol Cell Biol *12*, 4412-4421.

Lai, E.C. (2002). Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nat Genet *30*, 363-364.

Lai, W.S., Parker, J.S., Grissom, S.F., Stumpo, D.J., and Blackshear, P.J. (2006). Novel mRNA targets for tristetraprolin (TTP) identified by global analysis of stabilized transcripts in TTP-deficient fibroblasts. Mol Cell Biol *26*, 9196-9208.

Lam, F.H., Steger, D.J., and O'Shea, E.K. (2008). Chromatin decouples promoter threshold from dynamic range. Nature *453*, 246-250.

Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., *et al.* (2007). A mammalian microRNA expression atlas based on small RNA library sequencing. Cell *129*, 1401-1414.

Ledebur, H.C., and Parks, T.P. (1995). Transcriptional regulation of the intercellular adhesion molecule-1 gene by inflammatory cytokines in human endothelial cells. Essential roles of a variant NF-kappa B site and p65 homodimers. J Biol Chem *270*, 933-943.

Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798.

Lim, L.P., Lau, N.C., Garrett-Engele, P., Grimson, A., Schelter, J.M., Castle, J., Bartel, D.P., Linsley, P.S., and Johnson, J.M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature *433*, 769-773.

Liou, H.C., Nolan, G.P., Ghosh, S., Fujita, T., and Baltimore, D. (1992). The NF-kappa B p50 precursor, p105, contains an internal I kappa B-like inhibitor that preferentially inhibits p50. EMBO J *11*, 3003-3009.

Litvak, V., Ramsey, S.A., Rust, A.G., Zak, D.E., Kennedy, K.A., Lampano, A.E., Nykter, M., Shmulevich, I., and Aderem, A. (2009). Function of C/EBPdelta in a regulatory circuit that discriminates between transient and persistent TLR4-induced signals. Nat Immunol *10*, 437-443.

Liu, C.G., Calin, G.A., Meloon, B., Gamliel, N., Sevignani, C., Ferracin, M., Dumitru, C.D., Shimizu, M., Zupo, S., Dono, M., *et al.* (2004). An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc Natl Acad Sci U S A *101*, 9740-9744.

Maerkl, S.J., and Quake, S.R. (2007). A systems approach to measuring the binding energy landscapes of transcription factors. Science *315*, 233-237.

Mahtani, K.R., Brook, M., Dean, J.L., Sully, G., Saklatvala, J., and Clark, A.R. (2001). Mitogen-activated protein kinase p38 controls the expression and posttranslational modification of tristetraprolin, a regulator of tumor necrosis factor alpha mRNA stability. Mol Cell Biol *21*, 6461-6469.

Mark Ptashne, A.G. (2002). Genes & Signals. Cold Spring Harbor laboratory Press.

Marshak-Rothstein, A. (2006). Toll-like receptors in systemic autoimmune disease. Nat Rev Immunol *6*, 823-835.

Martone, R., Euskirchen, G., Bertone, P., Hartman, S., Royce, T.E., Luscombe, N.M., Rinn, J.L., Nelson, F.K., Miller, P., Gerstein, M., *et al.* (2003). Distribution of NF-kappaB-binding sites across human chromosome 22. Proc Natl Acad Sci U S A *100*, 12247-12252.

Mathys, S., Schroeder, T., Ellwart, J., Koszinowski, U.H., Messerle, M., and Just, U. (2003). Dendritic cells under influence of mouse cytomegalovirus have a physiologic dual role: to initiate and to restrict T cell activation. J Infect Dis *187*, 988-999.

Mayo, A.E., Setty, Y., Shavit, S., Zaslaver, A., and Alon, U. (2006). Plasticity of the cis-regulatory input function of a gene. PLoS Biol *4*, e45.

Medzhitov, R. (2008). Origin and physiological roles of inflammation. Nature *454*, 428-435.

Medzhitov, R., and Horng, T. (2009). Transcriptional control of the inflammatory response. Nat Rev Immunol *9*, 692-703.

Nathan, C. (2002). Points of control in inflammation. Nature *420*, 846-852.

Natoli, G. Maintaining cell identity through global control of genomic organization. Immunity *33*, 12-24.

Nau, G.J., Richmond, J.F., Schlesinger, A., Jennings, E.G., Lander, E.S., and Young, R.A. (2002). Human macrophage activation programs induced by bacterial pathogens. Proc Natl Acad Sci U S A *99*, 1503-1508.

Neish, A.S., Read, M.A., Thanos, D., Pine, R., Maniatis, T., and Collins, T. (1995). Endothelial interferon regulatory factor 1 cooperates with NF-kappa B as a transcriptional activator of vascular cell adhesion molecule 1. Mol Cell Biol *15*, 2558-2569.

Nenan, S., Boichot, E., Lagente, V., and Bertrand, C.P. (2005). Macrophage elastase (MMP-12): a pro-inflammatory mediator? Mem Inst Oswaldo Cruz *100 Suppl 1*, 167-172.

Nishikura, K. (2006). Editor meets silencer: crosstalk between RNA editing and RNA interference. Nat Rev Mol Cell Biol *7*, 919-931.

O'Connell, R.M., Taganov, K.D., Boldin, M.P., Cheng, G., and Baltimore, D. (2007). MicroRNA-155 is induced during the macrophage inflammatory response. Proc Natl Acad Sci U S A *104*, 1604-1609.

Oganesyan, G., Saha, S.K., Guo, B., He, J.Q., Shahangian, A., Zarnegar, B., Perry, A., and Cheng, G. (2006). Critical role of TRAF3 in the Toll-like receptor-dependent and -independent antiviral response. Nature *439*, 208-211.

Ogawa, S., Lozach, J., Benner, C., Pascual, G., Tangirala, R.K., Westin, S., Hoffmann, A., Subramaniam, S., David, M., Rosenfeld, M.G., *et al.* (2005). Molecular determinants of crosstalk between nuclear receptors and toll-like receptors. Cell *122*, 707-721.

Ogawa, S., Lozach, J., Jepsen, K., Sawka-Verhelle, D., Perissi, V., Sasik, R., Rose, D.W., Johnson, R.S., Rosenfeld, M.G., and Glass, C.K. (2004). A nuclear receptor corepressor transcriptional checkpoint controlling activator protein 1-dependent gene networks required for macrophage activation. Proc Natl Acad Sci U S A *101*, 14461-14466.

Panne, D., Maniatis, T., and Harrison, S.C. (2007). An atomic model of the interferon-beta enhanceosome. Cell *129*, 1111-1123.

Parekh, B.S., and Maniatis, T. (1999). Virus infection leads to localized hyperacetylation of histones H3 and H4 at the IFN-beta promoter. Mol Cell *3*, 125-129.

Park, C., Lecomte, M.J., and Schindler, C. (1999). Murine Stat2 is uncharacteristically divergent. Nucleic Acids Res *27*, 4191-4199.

Patterson, J.B., and Samuel, C.E. (1995). Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase. Mol Cell Biol *15*, 5376-5388.

Pedersen, I.M., Cheng, G., Wieland, S., Volinia, S., Croce, C.M., Chisari, F.V., and David, M. (2007). Interferon modulation of cellular microRNAs as an antiviral mechanism. Nature *449*, 919-922.

Perissi, V., Aggarwal, A., Glass, C.K., Rose, D.W., and Rosenfeld, M.G. (2004). A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors. Cell *116*, 511-526.

Plaksin, D., Baeuerle, P.A., and Eisenbach, L. (1993). KBF1 (p50 NF-kappa B homodimer) acts as a repressor of H-2Kb gene expression in metastatic tumor cells. J Exp Med *177*, 1651-1662.

Ptashne, M., and Gann, A. (2002). Genes & Signals. Cold Spring Harbor laboratory Press.

Ragoussis, J., Field, S., and Udalova, I.A. (2006). Quantitative profiling of protein-DNA binding on microarrays. Methods Mol Biol *338*, 261-280.

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.

Ramsey, S.A., Klemm, S.L., Zak, D.E., Kennedy, K.A., Thorsson, V., Li, B., Gilchrist, M., Gold, E.S., Johnson, C.D., Litvak, V., *et al.* (2008). Uncovering a macrophage transcriptional program by integrating evidence from motif scanning and expression dynamics. PLoS Comput Biol *4*, e1000021.

Roach, J.C., Smith, K.D., Strobe, K.L., Nissen, S.M., Haudenschild, C.D., Zhou, D., Vasicek, T.J., Held, G.A., Stolovitzky, G.A., Hood, L.E., *et al.* (2007). Transcription factor expression in lipopolysaccharide-activated peripheral-blood-derived mononuclear cells. Proc Natl Acad Sci U S A *104*, 16245-16250.

Rodriguez, A., Vigorito, E., Clare, S., Warren, M.V., Couttet, P., Soond, D.R., van Dongen, S., Grocock, R.J., Das, P.P., Miska, E.A., *et al.* (2007). Requirement of bic/microRNA-155 for normal immune function. Science *316*, 608-611.

Sandler, H., and Stoecklin, G. (2008). Control of mRNA decay by phosphorylation of tristetraprolin. Biochem Soc Trans *36*, 491-496.

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

Scadden, A.D. (2005). The RISC subunit Tudor-SN binds to hyper-edited double-stranded RNA and promotes its cleavage. Nat Struct Mol Biol *12*, 489-496.

Schreiber, J., Jenner, R.G., Murray, H.L., Gerber, G.K., Gifford, D.K., and Young, R.A. (2006). Coordinated binding of NF-kappaB family members in the response of human cells to lipopolysaccharide. Proc Natl Acad Sci U S A *103*, 5899-5904.

Segal, E., Friedman, N., Koller, D., and Regev, A. (2004). A module map showing conditional activity of expression modules in cancer. Nat Genet *36*, 1090-1098.

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

Sharma, S., tenOever, B.R., Grandvaux, N., Zhou, G.P., Lin, R., and Hiscott, J. (2003). Triggering the interferon antiviral response through an IKK-related pathway. Science *300*, 1148-1151.

Stark, A., Brennecke, J., Bushati, N., Russell, R.B., and Cohen, S.M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. Cell *123*, 1133-1146.

Stefani, G., and Slack, F.J. (2008). Small non-coding RNAs in animal development. Nat Rev Mol Cell Biol *9*, 219-230.

Stoecklin, G., Tenenbaum, S.A., Mayo, T., Chittur, S.V., George, A.D., Baroni, T.E., Blackshear, P.J., and Anderson, P. (2008). Genome-wide analysis identifies interleukin-10 mRNA as target of tristetraprolin. J Biol Chem *283*, 11689-11699.

Swanson, J.A., Lee, M., and Knapp, P.E. (1991). Cellular dimensions affecting the nucleocytoplasmic volume ratio. J Cell Biol *115*, 941-948.

Taganov, K.D., Boldin, M.P., Chang, K.J., and Baltimore, D. (2006). NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. Proc Natl Acad Sci U S A *103*, 12481-12486.

Takeuchi, O., and Akira, S. Pattern recognition receptors and inflammation. Cell *140*, 805-820.

Takeuchi, O., and Akira, S. (2010). Pattern recognition receptors and inflammation. Cell *140*, 805-820.

Tallquist, M., and Kazlauskas, A. (2004). PDGF signaling in cells and mice. Cytokine Growth Factor Rev *15*, 205-213.

Ten, R.M., Paya, C.V., Israel, N., Le Bail, O., Mattei, M.G., Virelizier, J.L., Kourilsky, P., and Israel, A. (1992). The characterization of the promoter of the gene encoding the p50 subunit of NF-kappa B indicates that it participates in its own regulation. EMBO J *11*, 195-203.

Thai, T.H., Calado, D.P., Casola, S., Ansel, K.M., Xiao, C., Xue, Y., Murphy, A., Frendewey, D., Valenzuela, D., Kutok, J.L., *et al.* (2007). Regulation of the germinal center response by microRNA-155. Science *316*, 604-608.

Thanos, D., and Maniatis, T. (1995). Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. Cell *83*, 1091-1100.

Udalova, I.A., Richardson, A., Denys, A., Smith, C., Ackerman, H., Foxwell, B., and Kwiatkowski, D. (2000). Functional consequences of a polymorphism affecting NF-kappaB p50-p50 binding to the TNF promoter region. Mol Cell Biol *20*, 9113-9119.

Valencia-Sanchez, M.A., Liu, J., Hannon, G.J., and Parker, R. (2006). Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev *20*, 515-524.

Wang, Q., Miyakoda, M., Yang, W., Khillan, J., Stachura, D.L., Weiss, M.J., and Nishikura, K. (2004). Stress-induced apoptosis associated with null mutation of ADAR1 RNA editing deaminase gene. J Biol Chem *279*, 4952-4961.

Warren, C.L., Kratochvil, N.C., Hauschild, K.E., Foister, S., Brezinski, M.L., Dervan, P.B., Phillips, G.N., Jr., and Ansari, A.Z. (2006). Defining the sequence-recognition profile of DNA-binding molecules. Proc Natl Acad Sci U S A *103*, 867-872.

Waterfield, M.R., Zhang, M., Norman, L.P., and Sun, S.C. (2003). NF-kappaB1/p105 regulates lipopolysaccharide-stimulated MAP kinase signaling by governing the stability and function of the Tpl2 kinase. Mol Cell *11*, 685-694.

Wei, L., Sandbulte, M.R., Thomas, P.G., Webby, R.J., Homayouni, R., and Pfeffer, L.M. (2006). NFkappaB negatively regulates interferon-induced gene expression and anti-influenza activity. J Biol Chem *281*, 11678-11684.

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

Wienholds, E., and Plasterk, R.H. (2005). MicroRNA function in animal development. FEBS Lett *579*, 5911-5922.

Yang, J.H., Luo, X., Nie, Y., Su, Y., Zhao, Q., Kabir, K., Zhang, D., and Rabinovici, R. (2003). Widespread inosine-containing mRNA in lymphocytes regulated by ADAR1 in response to inflammation. Immunology *109*, 15-23.

Yang, W., Chendrimada, T.P., Wang, Q., Higuchi, M., Seeburg, P.H., Shiekhattar, R., and Nishikura, K. (2006). Modulation of microRNA processing and expression through RNA editing by ADAR deaminases. Nat Struct Mol Biol *13*, 13-21.

Zhu, X., Hart, R., Chang, M.S., Kim, J.W., Lee, S.Y., Cao, Y.A., Mock, D., Ke, E., Saunders, B., Alexander, A., *et al.* (2004). Analysis of the major patterns of B cell gene expression changes in response to short-term stimulation with 33 single ligands. J Immunol *173*, 7141-7149.