SPECIFICITY OF TRANSCRIPTION ACTIVATION BY NF-κB SUBUNITS

Thesis by

Thomas Hin-Chai Leung

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2005

(Defended 9 September 2004)

ACKNOWLEDGMENTS

There are so many people that have been instrumental in my foray into science. I'd first like to thank my thesis advisor, David Baltimore. My experience in your lab has been marvelous, and I couldn't have asked for a better advisor. I thank you for your steadfast resolve in my growth as a scientist and for championing my work. I venerate you for your thoughtfulness, your uncanny ability to codify people's work to their core question, and most of all your worldliness. You have been a wonderful role model.

My thesis committee, Professors David Chan, Ellen Rothenberg, and Paul Sternberg, have been wonderful with their time, support, and ideas. I am grateful for every comment and suggestion that improved my project.

I owe a great deal of gratitude to the entire Baltimore Lab. In particular, I would like to single out Alexander Hoffmann. Working in an intimidating environment, you are a magnanimous mentor who gave your time to guide a tenderfoot through the rigors of science and life. You have influenced me in countless ways. I hope to be as good of a mentor as you were to me.

Eric Brown, Mark Boldin, Huatao Guo, Carlos Lois, Wange Lu, Mollie Meffert, and Xiao-Feng Qin all generously gave scientific and technical advice. In addition to giving technical help, Joel Pomerantz challenged me to think and to discover science by myself. I will always remember that. Elisha Gaston, you've provided tremendous help to all my projects. I hope you're getting a good feel for science. Lili Yang, we made it! Jeff Wiezorek, thanks for teaching me to enjoy the experience. Markus Covert, I enjoy all of our interactions and thanks for helping to proofread all my writings. Finally, I thank Elissa Denny for always having a smile ready and an offer to help out on anything. I thank the Alberola lab (especially Jose Alberola-Ila, Harry Green, Micheline Laurent, and Eric Tse) for technical advice and thoughtful interactions.

Special thanks go to my advisor at UCLA, Steve Smale, for supporting my decision to pursue my PhD at Caltech. Also to Dr. Stan Korenman, Alice Del Rosario, Professor Paul Patterson and

Elizabeth Ayala for seamlessly running the UCLA/Caltech Medical Scientist Training Program and for taking a chance on me.

Finally, there aren't enough words to describe my gratitude to my parents, Lawrence and Jean, and my sister, Rita. All of my success is a direct result of your influence and unconditional love.

ABSTRACT

The transcription factor NF-κB is a regulator of a wide variety of processes including inflammation, innate and adaptive immunity, apoptosis, and learning. How can one factor be accurately involved in so many different processes and generate without exception precise and appropriate responses? Four members of the NF-κB transcription factor family are involved in gene activation, and they hetero- or homodimerize with each other to bind DNA. This allows for many different potential combinations of NF-κB dimers. To test whether NF-κB-dependent genes require specific NF-κB family members for gene activation, cell lines were generated lacking in individual and multiple NF-κB proteins. Using TNFα as an inducer, a panel of endogenous NF-κB responsive genes showed a wide range of subunit specificities. Given that the NF-κB consensus binding site sequence is very broad and that crystal structures of NF-κB have not identified enough dimer-specific DNA-binding contacts to rationalize specific NF-κB binding sites, κB sites were compared from a single gene to another and no direct correlation was found between κb site sequence and κB family member requirements. However when interspecies comparisons were made of the same gene, a remarkable constancy of the κB site sequence was found, which suggested that individual sites have important functional characteristics. To test this theory, a novel lentiviral system was created that incorporated regulatory sequences into cellular DNA. Then by simply swapping sites between κB-dependent genes, NF-κB dimer specificity of the promoters was altered and revealed that two κB sites can function together as a module to regulate gene activation. Further, although the sequence of the κB site is important for determining κB family member specificity, rather than determining the ability of a particular dimer to bind effectively, the sequence affects which co-activators will form productive interactions with the bound κB dimer. My findings suggest that a particular DNA-binding site may impart a specific configuration to bound transcription factors that specifies the requirement for particular co-activators. Taken together, I have taken the first steps to dissecting how the promoter code influences individual NF-κB family members to function on NF-κB responsive genes and to regulate gene expression.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

Transcription Specificity

In mammalian cells, signaling molecules such as transcription factors are often encoded into gene families where members have distinct and overlapping functions. However, little is known about how individual family members within a particular transcription factor family operate. Study of this aspect of transcription control is often confounded by the recognition of a single regulatory sequence by multiple members of a transcription factor family. For example, *in vitro* studies have shown that all members of the basic helix-loop-helix (bHLH) leucine zipper transcription factor family recognize the same DNA-binding motif. While this result may be interpreted as functional redundancy among bHLH transcription factors, *in vivo* studies have demonstrated discrimination among particular family members at a particular site (Mai S., 1995). The fact that increasingly complex genomes contain a greater number of gene family members rather than novel families (Lander et al., 2001) suggests that diversity within a gene family may provide for specificity and versatility in the regulation of cellular responses. Characterizing the specific functions of individual members of transcription factor gene families contributes to our understanding of gene regulatory networks and to the development of specific therapeutic strategies.

The best understood example of transcription specificity is the nuclear hormone receptor (NHR) family, where individual members within the family have different intrinsic *in vivo* DNA-binding specificities. NHR family members bind two tandem cognate DNA half-sites, and biochemical studies have shown that the spacing between the two half-sites determines which NHR family members bind and function (Rastinejad, 2001). For example, thyroid hormone receptor binds to

DNA when four nucleotides separate the half-sites, whereas retinoic acid hormone receptor binds when the half-sites are separated by two nucleotides. This principle implies that adding or subtracting a number of nucleotides between the two DNA half-sites can alter a gene's responsiveness to a particular nuclear hormone receptor. Taken together, this evidence shows that the sequence of the NHR DNA-binding sites determines which family members bind and function. Recent structural studies with the glucocorticoid receptor (GR), a member of the NHR family, demonstrated that the DNA-binding domain of GR binds its cognate site in a variety of conformations depending on the sequence of the site (Lefstin and Yamamoto, 1998). The bound conformation served to present different interfaces for the recruitment of specific co-factors and to help determine whether the bound factor will activate or repress transcription. Further work has shown which GR surfaces are required in a gene-specific manner (Rogatsky et al., 2002; Rogatsky et al., 2003). Similar findings have also been reported for the POU family of proteins (Phillips and Luisi, 2000).

However, for many other transcription factor families, such as the C/EBP (Ramji and Foka, 2002), E2F (Trimarchi and Lees, 2002), AP-1 (Chinenov and Kerppola, 2001; Mechta-Grigoriou et al., 2001), and NF-κB families, marked DNA-binding specificity rules have not been identified. Indeed, the *raison d'être* of multiple members within signal transduction protein families may not be molecular interaction specificities but could be found in their own differential regulation.

Nuclear Factor kappa B (NF-κB)

NF-κB is a dimeric protein that mediates cellular responses to a wide variety of stimuli including TNFα, LPS, IL-1, and ultraviolet light (Ghosh et al., 1998; Karin and Ben-Neriah, 2000; Karin et

al., 2002). When our lab discovered this factor in 1986, NF-κB was thought to be involved in the control of the κ immunoglobulin locus in B cells (thus the name Nuclear Factor for κ in B cells). This assumption was quickly falsified when NF-κB was found to be widely expressed in many cell types and to be activated by a variety of stimuli. The list of systems that require NF-κB to function correctly has grown to encompass a large portion of biology, including inflammatory responses, immune system development, apoptosis, learning in the brain, and bone development. In resting cells, NF-κB is held inactive via associations with inhibitory proteins called inhibitors of NF-κB (IκB). When the cell is stimulated, IκB proteins are phosphorylated, ubiquitinated, and degraded, allowing NF-κB to bind DNA and activate appropriate target genes.

In response to inflammatory stimuli, four members of the NF-κB family are involved in gene activation: p50, p52, p65, and cRel. They are the mature gene products from four genes: nfκB1, nfκB2, relA, and cRel respectively, and they homo- or heterodimerize with one another to bind DNA. The cognate DNA-binding element, the κB-site, is found in the promoters of NF-κB responsive genes but displays a remarkably loose consensus, often described as $G_{-5}G_{-4}G_{-3}R_{-2}N_{-1}N_0Y_{+1}Y_{+1}C_{+2}C_{+3}$ (Chen et al., 1998a). The heterogeneity of κB-sites has been thought to confer specificity of regulation mediated by differential affinities of NF-κB isoforms. Purified recombinant RHD protein dimers could indeed be shown to select differential optimal DNA-binding motifs from a pool of random oligonucleotides (Kunsch et al., 1992). Similarly, one study with transfected promoter-reporter constructs showed that multimerized κB-sites from one of three different promoters (Igκ, MHC H2, and IFN-ß) exhibited differential responses when co-transfected with p50 and/or p65 expression plasmids (Fujita et al., 1992).

Structural studies have also not revealed evidence for highly specific interactions between NF-κB and its cognate DNA-binding sites. Crystal structures have shown that p50 homodimers often interact specifically with 5'-$G_{-5}G_{-4}G_{-3}$, while p65 homodimers interact specifically with 5'-$G_{-5}G_{-4}$. X-ray structures of p50:p65 heterodimers demonstrate that they can bind to a variety of κB site sequences (Berkowitz et al., 2002; Chen and Ghosh, 1999; Chen-Park et al., 2002; Escalante et al., 2002). Taken together, structural studies have not identified enough dimer-specific DNA-binding contacts to rationalize specific NF-κB binding sites and underscore the remarkable permissiveness in NF-κB-DNA interactions.

Despite the difficulty in discerning specificity at the level of DNA-protein interactions, mice deficient in a single RHD protein show specific phenotypes. *rela*$^{-/-}$ animals are embryonic lethal with massive hepatocyte apoptosis (Beg et al., 1995), while *crel*$^{-/-}$ mice develop normally but have defects in lymphocyte proliferation (Kontgen et al., 1995). In part, family member-specific functions can be inferred by tissue specific expression. While p65 appears to be ubiquitous, cRel is expressed constitutively in mature monocytic and lymphocytic lineages (Liou et al., 1994) and induced in other cell types, such as fibroblasts (Grumont and Gerondakis, 1990). In addition, there are differences in the molecular characteristics of RHD family members: transfection studies have identified activation domains in p65 and cRel, but not in p50 and p52 (Chen et al., 1998a). DNA-binding assays were used to identify dimerization rules: p50 (and presumably p52) can partner with all RHD family members including themselves, while RelB and cRel do not appear to engage in homotypic interactions. RelB, furthermore, does not appear to dimerize with p65 or cRel. Little is known about how the results from *in vitro* biochemical characterizations of RHD proteins can

serve to explain physiologically relevant functional specificity in gene regulation, because selectivity of endogenous κB-site containing promoters remains largely unexplored. In fact, in the case of the IL-12 promoter, results from cell-free or extra-chromosomal templates are shown not to recapitulate endogenous promoter specificity (Sanjabi et al., 2000).

Genetic approaches are suitable to address questions of *in vivo* mechanistic specificity, but must take into account the possibility of interdependent regulation of factors in mammalian signaling networks; resulting compensatory mechanisms among functionally related molecules may lend increased robustness to signaling systems and affect the phenotype of genetic deletions.

Overview of Thesis

The overall question driving force of this thesis is "how does one transcription factor control so many different processes without generating inappropriate responses in particular settings?" While NF-κB has been shown to be an important pillar in modulating cellular behavior in response to varied needs of organismal life, we know very little mechanistically of its functions. Understanding how NF-κB is regulated, how it interacts with other signaling pathways or molecules, and how it behaves on gene promoters will aid in understanding disease processes and aid in future drug design. Right now, the first generation drugs that target NF-κB activities are being employed for osteoporosis, cancer, arthritis and sepsis. Deeper knowledge of this key system will facilitate the creation of more specific drugs and therapies.

With this framework established, we decided to look at the DNA level for answers. Before I begin on my own work, Chapter 2 is an optional read for someone interested in the history of the field.

There are plenty of good books on the topic, but I felt it is important for me to write about it and to show how my work is placed in the general context of the field. Chapter 3 describes how I collaborated with Alexander Hoffmann to create a genetic system to address whether NF-κB-dependent genes require particular κB family members for activation. Using 3T3 cell lines lacking individual and multiple NF-κB proteins and TNFα as an inducer, I found that a panel of endogenous NF-κB-dependent genes showed a wide range of subunit requirements. After establishing that transcription specificity occurs in the NF-κB family, I wanted to address how specificity is determined and maintained. Chapter 4 summarizes our studies where I found that the sequence of the κB binding site was important for determining family member specificity. However rather than determining the ability of a particular dimer to bind effectively, the sequence affects which co-activators will form productive interactions with the bound κB dimer. Finally in Chapter 5, future studies are considered.

References

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

Berkowitz, B., Huang, D. B., Chen-Park, F. E., Sigler, P. B., and Ghosh, G. (2002). The x-ray crystal structure of the NF-kappa B p50.p65 heterodimer bound to the interferon beta -kappa B site. *J. Biol. Chem. 277*, 24694-24700.

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

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

Chen-Park, F. E., Huang, D. B., Noro, B., Thanos, D., and Ghosh, G. (2002). The kappa B DNA sequence from the HIV long terminal repeat functions as an allosteric regulator of HIV transcription. *J. Biol. Chem. 277*, 24701-24708.

Chinenov, Y., and Kerppola, T. K. (2001). Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. *Oncogene 20*, 2438-2452.

Escalante, C. R., Shen, L., Thanos, D., and Aggarwal, A. K. (2002). Structure of NF-kappaB p50/p65 heterodimer bound to the PRDII DNA element from the interferon-beta promoter. *Structure* (Camb) *10*, 383-391.

Fujita, T., Nolan, G. P., Ghosh, S., and Baltimore, D. (1992). Independent modes of transcriptional activation by the p50 and p65 subunits of NF-kappa B. *Genes Dev. 6*, 775-787.

Ghosh, S., May, M. J., and Kopp, E. B. (1998). NF-kB and Rel proteins: Evolutionarily conserved mediators of immune responses. *Annu. Rev. Immunol*. *16*, 225-260.

Grumont, R. J., and Gerondakis, S. (1990). Murine c-rel transcription is rapidly induced in T-cells and fibroblasts by mitogenic agents and the phorbol ester 12-O-tetradecanoylphorbol-13-acetate. *Cell Growth Differ. 1*, 345-350.

Karin, M., and Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. *Annu. Rev. Immunol. 18*, 621-663.

Karin, M., Cao, Y., Greten, F. R., and Li, Z. W. (2002). NF-kappaB in cancer: from innocent bystander to major culprit. *Nat. Rev. Cancer 2*, 301-310.

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

Kunsch, C., Ruben, S. M., and Rosen, C. A. (1992). Selection of optimal 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.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature 409*, 860-921.

Lefstin, J. A., and Yamamoto, K. R. (1998). Allosteric effects of DNA on transcriptional regulators. *Nature 392*, 885-888.

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

Mai S., a. M. I. L. (1995). The c-myc Protein Represses the Lambda5 and Tdt Initiators. *Nucleic Acids Research 23*, 1-9.

Mechta-Grigoriou, F., Gerald, D., and Yaniv, M. (2001). The mammalian Jun proteins: redundancy and specificity. *Oncogene 20*, 2378-2389.

Phillips, K., and Luisi, B. (2000). The virtuoso of versatility: POU proteins that flex to fit. *J. Mol. Biol. 302*, 1023-1039.

Ramji, D. P., and Foka, P. (2002). CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem. J. 365*, 561-575.

Rastinejad, F. (2001). Retinoid X receptor and its partners in the nuclear receptor family. *Curr. Opin. Struct. Biol. 11*, 33-38.

Rogatsky, I., Luecke, H. F., Leitman, D. C., and Yamamoto, K. R. (2002). Alternate surfaces of transcriptional coregulator GRIP1 function in different glucocorticoid receptor activation and repression contexts. *Proc. Natl. Acad. Sci.* USA *99*, 16701-16706.

Rogatsky, I., Wang, J. C., Derynck, M. K., Nonaka, D. F., Khodabakhsh, D. B., Haqq, C. M., Darimont, B. D., Garabedian, M. J., and Yamamoto, K. R. (2003). Target-specific utilization of transcriptional regulatory surfaces by the glucocorticoid receptor. *Proc. Natl. Acad. Sci*. USA *100*, 13845-13850.

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.* USA *97*, 12705-12710.

Trimarchi, J. M., and Lees, J. A. (2002). Sibling rivalry in the E2F family. *Nat. Rev. Mol. Cell. Biol. 3*, 11-20.

CHAPTER 2: A BRIEF HISTORY OF GENE REGULATION

Regulation of gene transcription is an important mechanism conserved from bacteria to humans. In bacteria, gene control predominantly serves to adjust a cell's response to changes in the environment. In multicellular organisms, changes in the environment also alter gene expression. However, the most complex example of multicellular gene control is embryogenesis, where multiple cell types work flawlessly together to execute multiple genetic programs.

The birth of modern biochemistry undoubtedly facilitated initial discoveries with respect to transcription control. Early biochemical studies demonstrated that the same enzymes and pathways in yeast and mammalian cells were responsible for the conversion of glycogen to lactic acid. This and other discoveries made it clear that core biological mechanisms and molecules are preserved over evolutionary times. This has also been proven to hold true for transcription control. The concepts of promoters, allosteric regulation, and cooperative binding originally identified in bacteria by Monod, Jacob, and Lwoff are also important in yeast and higher eukaryotes. Eukaryotic cells are more complex than bacteria, and studies by Ptashne and Chambon have enumerated on these original findings and added the ideas of modular transcription factors, nucleosome, and enhancers. More recent studies in higher eukaryotes by Maniatis and Chambon led to the concept of the enhanceosome. The field of gene transcription has had a rich history filled with people of great stature, and this chapter highlights the important advances in transcription regulation with a particular focus on activators and repressors. Information concerning the history and identification of the basal transcriptional machinery itself (for example TBP, TAFs, mediator

complex, nucleosome modifiers) has been consciously omitted, because my thesis work mainly concerns activators and the molecules they recruit.

Lessons from Bacteria: *lacZ* and the lambda repressor

The enzyme ß-galactosidase, the product of the *lacZ*, is the first enzyme to act in bacterial metabolism of lactose. The gene is fully activated when bacteria are grown in glucose-deficient, lactose-containing bacterial medium. Glucose is an efficient energy source for bacteria and normally overrides lactose metabolism. How does this regulation operate? Two DNA-binding proteins control *lacZ* regulation: CAP (Catabolite Activator Protein) and the *lac*-Repressor. In the absence of lactose and irrespective of the absence or presence of glucose, *lac*-Repressor binds its DNA-binding-site located in the operator region within the *lac* operon. The presence of *lac*-Repressor prevents RNA polymerase from binding to the *lac* operon. In the absence of glucose, the CAP protein (also known as CRP or cAMP Receptor Protein) binds the *lacZ* operator and recruits RNA polymerase to the gene. When glucose and lactose are both present, *lac*-Repressor is removed, but the absence of CAP prevents direct RNA polymerase recruitment and transcription is not induced. Levels of *lacZ* transcript are marginally higher than resting levels, because RNA polymerase can randomly bind its exposed operator site. When both glucose and lactose are absent, CAP and *lac*-Repressor are bound to their respective sites on the *lac* operon; *lac*-Repressor still prevents RNA polymerase binding and transcription is repressed.

CAP and *lac*-Repressor are activated by their physiologic signals in an allosteric fashion. In the absence of glucose, bacterial cells synthesize more cAMP protein. cAMP binds and changes CAP's (hence the original cAMP Receptor Protein or CRP designation) conformation, such that

CAP binds DNA more efficiently. *lac*-Repressor undergoes a similar method of regulation. In the presence of lactose, allolactose (a product of lactose breakdown) binds and changes *lac*-Repressor's conformation such that *lac*-Repressor no longer binds DNA.

It is remarkable that Monod and Jacob were able to piece together the principles of this switch using genetics and bacterial conjugation. Monod made the observation that the growth curve of *E. coli* differed depending on the presence of different carbohydrate pairs in the medium. He believed that this was due to "enzyme adaptation" and devoted his life to the study of this phenomenon. He established the ß-galactosidase system, understood that bacterial growth in that system was dependent on the synthesis of an astable enzyme, and isolated mutants which revealed that several key proteins were dependent on different genetic elements yet subjected to the same induction determinism and constitutive mutations. Concurrently, Jacob had been studying bacterial conjugation and discovered that conjugation passed along the male chromosome without the male cytoplasm into the female bacterium. Further, he developed a system to monitor kinetics of conjugation (limiting how much of the chromosome is transferred by using a kitchen blender) and another system of making "diploid" bacterium. Jacob and Monod could now perform what we call now–reconstitution or rescue experiments. Mutant bacteria could be injected with wild-type copies of genes to see if function is "rescued." For example, bacteria that produced ß-galactosidase constitutively were reconstituted with the genes for CAP or *lacZ* Repressor. From the CAP protein reconstitution, induction of the introduced gene occurred at the maximum rate from the beginning (this finding led to the basis of the messenger RNA theory by Jacob). The *lac*-Repressor reconstitution showed that the gene was dominant to the constitutive mutation but that this

dominance occurred in a slow plodding fashion. This experiment was the basis for the "repressor" theory.

Other work demonstrated CAP's function to recruit RNA polymerase to the *lacZ* promoter. DNAse1 footprinting experiments showed that RNA polymerase bound to the *lacZ* promoter much more reliably in the presence of CAP. CAP mutants were also identified that could bind the *lacZ* promoter, but RNA polymerase recruitment was absent and transcription was inhibited. The region identified by the CAP mutant experiments interacted with the □-subunit of bacterial RNA polymerase. □-subunit RNA polymerase mutants were identified that could not activate *lacZ* in the presence of wildtype CAP protein.

However, the most compelling proof of regulated recruitment was provided by three activator-bypass experiments or experiments designed to bring RNA polymerase to the promoter in an alternative fashion. The first used synthetic heterologous protein:protein interactions. The DNA-binding-domain of CAP was fused to a synthetic interacting protein. The synthetic interacting protein's partner was fused to RNA polymerase. Expression of both mutant proteins activated lacz transcription. Presumably, the synthetic interacting protein directed itself to the *lacZ* promoter via CAP's DNA-binding-domain and recruited its partner to the same location. The presence of RNA polymerase stimulated transcription of the reporter gene. The second method involved fusing CAP's DNA-binding-domain directly to RNA polymerase to entirely bypass the requirement for CAP. Expression of this mutant polymerase activated transcription of *lacZ* in the presence or absence of CAP protein. The third experiment simply increased the amount of polymerase protein in the cell. If CAP's main function was to recruit RNA polymerase to particular genes, increasing

the concentration of free RNA polymerase would bypass the need for CAP. Together, these experiments provide compelling evidence that CAP's main function is to recruit RNA polymerase to the *lacZ* operon.

In recent years, attention has been put on a molecular view of *lacZ* regulation. Crystal structures of *lac*-Repressor bound to operator and inducer have identified the regions within *lac*-Repressor responsible for interaction. Recently, Kalodimos and colleagues solved the remaining piece of the puzzle, a crystal structure of *lac*-Repressor bound to nonspecific DNA that has proved most insightful. Interactions between transcription factors and nonspecific DNA are important in several ways. 1. Nonspecific sites participate in the regulation of physiological function, because transcription factors are often complexed with them *in vivo* when not bound at their regulatory sites. 2. When nonspecific DNA-binding is allowed, it has been shown that the binding kinetics for transcription factors occur much faster than the kinetics based on simple diffusion. Different mechanisms have been suggested to reconcile this difference that all depend on "facilitated transfer." 3. Even though a transcription factor binds its own binding site with high affinity, it also binds nonspecific sites with appreciable and significant affinity. For example, *lac*-Repressor binds the operator at 10^8-fold stronger than binding at a nonspecific site, but there are 10^7 potential nonspecific binding sites in the *E. coli* genome.

The crystal structure showed that *lac*-Repressor bound to nonspecific DNA is tilted by some 25 angstroms compared to the repressor-operator complex. As a result, the residues identified by the specific repressor-operator complex to interact directly with base pairs in the major groove of the operator are shifted. Instead, the residues participate in hydrogen bonds or electrostatic

interactions with the phosphate groups of nonspecific DNA. Mutating these residues lowers the affinity of repressor for operator some 100-fold and also lowers affinity for nonspecific DNA by 10-fold. These results highlight the dual-roles for the recognition region of *lac*-Repressor. The region confers specificity when bound to the operator; at the same time, the region stabilizes *lac*-Repressor when bound to nonspecific DNA. Further, when *lac*-Repressor binds to operator, a hinge region that is speculated to play a major role in induction, transitions from a disordered state into an alpha helix. When *lac*-Repressor binds nonspecific DNA, the hinge region remains in a disordered state. Finally, conformation studies performed in biologically relevant time showed that *lac*-Repressor is extremely flexible, open and variable on nonspecific DNA but adopts a more rigid structure when bound to operator. In summary, this is all evidence for a structural switch within the *lac*-Repressor that directs the binding mode depending on the sequence of DNA.

References: (Berg et al., 1981; Jacob, 1972; Kalodimos et al., 2004; Lewis et al., 1996; Monod, 1972; Ptashne and Gann, 2002)

Lambda Phage

Down the hall at the Institute Pasteur where the *lacZ* experiments were first being performed, Andre Lwoff and his group were studying lambda phage. They noticed that when infected *Escherichia coli* were irradiated with ultraviolet light for 45 minutes, bacteria stopped growing and roughly 90 minutes later they lysed, spewing a crop of viruses called phage (or bacteriophages or bacteria-eaters) into the medium. The newly released phage can then infect more bacteria and multiply within them. Growth of phage follows two forms: lysogenic and lytic. During lysogeny, lamba phage is integrated into the genome, and transcription of its own genome is kept to a

minimum. (Lwoff could passage lysogenically-infected bacteria over 19 generations without lysis.) During lytic growth, lambda phage is tremendously active, replicating furiously until endolysin is synthesized and the host cell is lysed releasing phage into the medium. Initially, Lwoff planned to study lysogeny and how the prophage interacts with the host machinery but with these findings induction became his focus.

The main principles learned from *lacZ* apply here, although the lambda switch is a bit more complex. Within the OR operon, a central operator region composed of three binding sites, O_R1, O_R2, and O_R3, plays the prominent role. During lysogeny, lamba repressor protein binds to the O_R1 and O_R2 regions. This forces RNA polymerase to bind O_R3 and to transcribe the cI gene. cI's gene product is the lambda repressor protein itself, therefore this is a feed-forward regulation loop (analogous to *lacZ*). Until the appropriate inductive signal occurs, this mechanism ensures that the lysogenic growth cycle of the lambda phage is maintained. UV light triggers lambda phage to enter its lytic growth cycle by activating RecA, a protein that promotes lambda repressor's autodigestion. The absence of lambda repressor on the operator reveals the O_R1 and O_R2 binding sites. O_R1 and O_R2 are inherently better binding sites than O_R3 for RNA polymerase. RNA polymerase binds to both sites and transcribes the cro genes in the opposite direction. Cro gene products facilitate the lytic cycle. By directing transcription away from cI, this elegant regulatory "switch" effectively shuts off the lysogenic growth cycle.

How does lysogeny ever become established? During the initial infection, cII protein is synthesized and binds a promoter site downstream of the cro genes. The presence of cII recruits RNA polymerase to the promoter and stimulates transcription in the opposite direction of cro and

actually promotes synthesis of lambda repressor. This allows lambda repressor levels to be established and entry into the lysogenic growth cycle. Whether phage establishes lysogeny or lytic growth upon initial infection is dependent on a host protease that attacks cII protein. If the cell is active and dividing, levels of the protease are high, cII protein is degraded, and phage enters the lytic growth cycle. If the cell is inactive, levels of the protease are low, cII protein is stable, and phage enters the lysogenic growth cycle.

Cooperativity plays a key role in the control of lambda repressor. A repressor dimer binding O_R1 helps another dimer bind the adjacent weak site O_R2. While this cooperativity does not continue to the O_R3 binding site, repressor may bind the O_R3 site and turn off transcription of its own gene. This autoregulatory behavior is consistent with the requirements of an efficient switch, because free unbound repressor would bind inappropriately during activating conditions. Because lambda repressor's affinity for the O_R3 binding site is lower than its affinity for O_R1 or O_R2, lambda repressor is able to regulate its own levels.

Current work

The activities of the repressor up to this point were identified by studying the OR region in isolation from the rest of the lambda phage chromosome. This approach was clearly fruitful, because investigators were able to identify individual components important in regulation. However, studies with the whole lambda chromosome have helped scientists to reformulate some of the original conclusions and demonstrate that the switch works even more efficiently than had originally been thought.

Recent studies have identified a second operator O_L positioned 2.4kB away from the O_R operator. O_L controls a different set of lambda phage genes and is similar to O_R in structure and function. O_L has three repressor binding sites O_L1, O_L2, and O_L3, with O_L1 also having the highest affinity. Repressor dimers bind O_L1 and O_L2 cooperatively, and repressor bound to O_L3 turns off transcription of the adjacent promoter (P_L). Previously, repressor was shown to form dimers and tetramers on O_R. On a DNA molecule containing O_R and O_L, repressor binds to O_R1, O_R2, O_L1, and O_L2 to form an octomer (or four dimers). The DNA between the two operators is looped out. "Long-range" cooperative binding between the four sites adds an additional level of interaction and increases repression at P_R and P_L a few fold. Another consequence of these interactions involves the autogenous control by repressor on itself. The "long-range" interaction between the two operators helps to facilitate repressor binding to O_R3. Initial studies with O_R3 in isolation rested on the assumption that *lac*-Repressor would bind <10% of the time, and *lac*-Repressor levels needed to increase 10-fold before it would bind O_R3 significantly. However, this long-range interaction actually increases the O_R3 occupancy rate to 60-70% and reduces resting repressor concentrations approximately 3-fold.

Why do the two operators function together?

If one considers a curve comparing the binding of repressor to the two operator sites as a function of repressor concentration, the curve would take a "sigmoid" or highly inflective shape. In other words, at a certain range small changes in repressor concentration would have dramatic effects on site occupancy. Cooperative-binding between the repressor dimers on OR plays a major role in defining the curve. Elimination of cooperativity between repressor monomers has the same effect as lowering overall concentration of repressor by 100-fold, even though lowering the concentration

of repressor 5-10-fold is enough to vacate the binding sites entirely. "Long-range" interactions only serve to make the curve steeper. In fact, in the absence of an inducing stimulus less than one in a million phage infected cells undergoes spontaneous change in growth cycles. In the presence of an inducing stimulus, virtually all infected cells undergo change. The lambda repressor is a remarkably efficient and effective switch.

References: (Dodd et al., 2001; Dodd et al., 2004; Jacob, 1972; Ptashne, 1986; Revet et al., 1999)

Lessons from Viruses: SV40

In eukaryotes, like bacteria, promoters direct where RNA polymerase binds and initiates transcription. However in eukaryotes, places where transcription factors bind and function are often located much farther away from the promoter that in bacterial genomes. These places can be thousands of base pairs away upstream or downstream from a transcription start site. The first enhancer discovered to affect eukaryotic gene transcription was in a 366 bp fragment from the SV40 (simian virus 40) genome. This enhancer normally serves to direct virus genome transcription. When a plasmid containing the SV40 fragment and any eukaryotic test promoter was transfected into cells, transcription was increased compared to the control. Further tests showed that orientation or distance away from the promoter had no effect on SV40's effectiveness. Analysis of the SV40 enhancer showed that it was made up of multiple elements that contribute to the overall effectiveness of the enhancer.

Lessons from Yeast: Gal4

Saccharomyces cerevisiae or Baker's yeast was quickly adopted as a system to study eukaryotic gene transcription. It replicates quickly (about 3-fold slower than bacteria), is easy to grow and to handle, and most importantly its genetic sequences can be moved or mutated in a convenient fashion. Eukaryotic gene expression is more complicated than bacterial gene expression because in eukaryotic cells, DNA is compacted by histone proteins into structures called nucleosomes. Further, eukaryotic RNA must be spliced and transported outside the nucleus for translation to occur. However, we'll stay focused on the mechanisms of transcription regulators by activators and repressors.

Studies have focused on a series of genes important in galactose metabolism. *gal1* is only activated when galactose is present in the medium and when glucose is absent. *S. cerevisiae* metabolizes glucose much more efficiently than it does galactose. When galactose is present, Gal4 binds the UAS (upper activating sequence) in the *gal1* promoter region and activates transcription by recruiting basal transcription machinery and RNA polymerase. If glucose is present in the medium and irrespective of the status of galactose, transcription is actively repressed, because the repressor Mig1 binds the *gal1* promoter and recruits the Tup1 repressive complex. As you can see, there are a lot of similarities between *gal1* and *lacZ*. Glucose metabolism dominates galactose or lactose metabolism by recruiting Mig1 or *lac*-Repressor respectively. The Gal4 and CAP activator proteins function by recruiting basal transcription machinery and RNA polymerase directly to the promoter.

Gal4 is normally associated with Gal80 and kept inactive. When a cell senses galactose, Gal3 activates and binds Gal80, changing Gal80's conformation. This change in conformation alters Gal80's association with Gal4 such that Gal4 can now bind DNA and recruit basal transcription machinery. In the absence of glucose, inactive Mig1 is phosphorylated and held in the cytoplasm. The presence of glucose degrades the kinase responsible for Mig1 phosphorylation. Unphosphorylated and active Mig1 enters the nucleus, binds DNA, and recruits the Tup1 repressive complex. This example re-emphasizes a popular mechanism among transcription regulation: inactive transcription regulators are usually excluded from the nucleus or have their activation domains camouflaged.

What separates these studies in yeast from those in bacteria was the recognition that transcription factors could be modular in nature. Gal4 is composed of two different and separable domains, a DNA-binding domain and an activation domain. Gal4 mutants that lack an activation or DNA-binding domain could not initiate transcription. When a mutant protein composed of the DNA-binding domain of the bacterial LexA protein attached to the Gal4-activating domain was introduced into cells, *gal1* containing a LexA-binding site in place of the Gal4-binding site could be activated. This experiment demonstrated that the activation-domain did not behave in a stereo-specific manner and could be attached to any DNA-binding domain. (And formed the basis for the popular yeast-two-hybrid technique.) To further prove this point, the Gal4-activating domain was fused directly to the Gal80 protein. Cells expressing this mutant protein expressed *gal1* irrespective of whether galactose was present in the medium.

How does the activation region work? Why does it behave equally well attached to any DNA-binding domain? The activation domain of Gal4 is "acidic" and composed of many negatively charged amino acids (like glutamic acid and aspartic acid) and hydrophobic residues. Mutations in the activation-domain that convert negatively charged residues to neutrally charged residues weaken the activation potential of the domain. Conversely, mutations in the activation-domain that convert neutrally charged residues to negatively charged residues strengthen the activation potential of the domain. The length of the activation domain is also important, with longer activation domains carrying stronger activation potentials. Finally, mutant proteins consisting of activation domains taken from bacterial proteins (e.g. LexA) fused with the Gal4 DNA-binding domain can still activate *gal1* transcription. Together, activation domains are transferable, and they function by recruiting basal transcription machinery to the promoter.

Activators that work by recruitment must bind DNA to serve their function. Thus if activators are significantly over-expressed in a cell, gene activation is inhibited. Excess unbound activators soak up available transcription machinery, leaving DNA-bound activators unable to perform their function. This phenomenon is called squelching and underscores the fact that transcription activators are highly regulated both in cellular location and activity.

References: (Ptashne and Gann, 2002)

Lessons from Higher Eukaryotes

The study of gene expression in higher eukaryotes is more complicated and more complex than in bacteria or viruses. It was presumed that more complex organisms needed more genes to handle the additional information. Comparative genome studies have refuted this notion. The *Caenorhabditis elegans* genome contains roughly 20,000 genes, while the *Drosophila* genome contains close to 14,000. However, *Drosophila* is clearly a more complex organism than the simple nematode worm. RNA splicing and DNA rearrangement are some potential ways for a smaller genome to provide added complexity. A third possibility is coordinate gene regulation programs where multiple factors are required for gene activation. This third option allows an individual protein to participate in different biological processes, because the protein functions only when appropriate binding partners are present. This combinatorial and modular approach of gene activation can add levels of complexity. In support of added regulation at the genetic level, there are roughly 300 transcription factors in yeast (counting both sequence-specific transcription factors and basal transcription machinery) compared to roughly 3000 transcription factors in humans. Since there are approximately 6000 genes in *Saccharomyces*, transcription factors make up 5% of the genome with a ratio of 1 transcription factor per 20 genes. The human genome contains roughly 30,000 genes, thus transcription factors make up 10% of the genome and there is a ratio of 1 transcription factor per 10 genes. A significant increase in the number of transcription factors may be all that is needed to account for the difference in complexity.

To further highlight the differences in complexity, a typical transcriptional unit from a unicellular eukaryote contains a simple core promoter, an upstream activator sequence, and a silencer element all within a region 100-200 bp away from the TATA box. In contrast, a typical higher eukaryotic

gene contains several enhancers that can be located 5'or 3' of the gene or even in introns. Each enhancer is likely to contain on average 10 binding sites for roughly three different sequence-specific transcription factors. The core promoter is still compact, but at least three different core promoter sequences have been identified to recruit TBP and RNA polymerase. Finally, the enhancers important for gene regulation have been found to be as far as 100 kb away from the transcription start site (see Igf2, Leighton et al., 1995; Webber et al., 1998).

This complicated organization of regulatory information must be central for allowing detailed control of gene expression, and the original hypothesis should now be revised. Levine and Tijian have recently proposed, "that physiological and behavioral complexity correlates with likely number of gene expression patterns exhibited during an animal's life cycle (Levine and Tjian, 2003)." Different angles have been used to dissect the mechanisms involved in implementing gene expression patterns. In particular, studies of the nuclear hormone receptor family have addressed how individual members within the family activate specific genes and identified the key players and mechanisms. On the other hand, studies of the *Drosophila* protein Dorsal have elegantly shown how a single protein coordinates dorsal-ventral patterning of an embryo by regulating numerous gene expression programs. Finally, studies with the IFNß and IL-2 genes have dissected how gene activation occurs on an individual promoter. These approaches all share the common goal of deciphering the "promoter code", where the regulation of a specific gene can be deduced by simple studying the promoter sequence.

Nuclear Hormone Receptor Family

The nuclear hormone receptor (NHR) transcription factor family has been the model for studying specificity within a transcription factor family. Lipophilic hormones regulate gene expression through this superfamily, which includes retinoic acid hormone receptor (RAR), thyroid hormone receptor (TR), and the peroxisome proliferation-activated receptor (PPAR). Within the superfamily, there exists a class of orphan receptors whose biological ligands have not been identified. Homologs of nuclear hormone receptors have been identified in insects, worms, and amphibians.

An important development in the NHR field was the identification and establishment of marked DNA binding specificity rules. Like all transcription factor families, transcription regulation depends on interactions between hormone/orphan receptors and the DNA sequence. Most non-steroid and orphan receptors (including RAR, TR, PPAR, vitamin D receptor (VDR) and nerve growth factor-induced B (NGFI-B)) recognize the consensus half-site sequence 5'-AGGTCA-3'. NHR family members form heterodimers with a common member of the NHR family, the retinoid X receptor (RXR). RXR binds DNA on its own duplicated half-site containing the same sequence. Importantly, interactions between receptors do not occur until the receptors are bound to DNA. Productive interactions between RXR and its partner are determined by the spacing between the two half-sites (Mader et al., 1993; Perlmann et al., 1993; Umesono et al., 1991; Zechel et al., 1994). The number of base pairs between two sites can range from one to five base pairs, and this rule has been generally referred to as the 1-5 rule. If four base pairs separate the two half-sites, TR-RXR heterodimers function; if two base pairs separate the two half-sites, PPAR-RXR heterodimers function. These rules have important implications for transcription specificity. By

simply changing the number of base pairs between two half-sites, signaling pathways that activate a particular gene can be interconverted. To add further complexity, the two half-sites can be distinguished between each other by their upstream and downstream locations. RXR can switch its polarity by binding specifically to an upstream or downstream site. For example, RAR-RXR dimers can bind to to half-sites separated by one or five bases, but they are organized in opposite polarities and functionally transform from an activator to a repressor respectively (Kurokawa et al., 1995).

Biochemical and biophysical studies have studied how RXR and its partners recognize spacing. Standard biochemical assays did not detect monomeric binding of RXR and individual receptors (Mader et al., 1993; Perlmann et al., 1993; Umesono et al., 1991; Zechel et al., 1994) which implied that protein-protein interactions must be used to achieve the desired binding affinity. Several crystal structures depicting heterodimers bound to DNA have been solved and show an asymmetrical placement of subunits in a head-to-tail fashion (Rastinejad et al., 1995; Rastinejad et al., 2000; Zhao et al., 2000). Each receptor interacted with the minor groove of the spacing DNA and formed protein-protein interactions with its partner. When the structure of a RXR-TR heterodimer bound to a four bp spacing was compared to a RXR-RAR heterodimer bound to a one bp spacing, there were significant differences in protein-protein interactions, the bending of DNA, and the conformational flexibility of the receptors. A detailed study of RXR binding affinity demonstrated that RXR in a RXR-RXR homodimer makes contact with 3 bp of the binding site, while RXR in a RXR-RAR heteromdimer makes contact with 6 bp of the binding site. Therefore RXR must adopt different conformations depending on its partner, and this flexibility may explain why RXR is the common partner. Taken together, biophysical studies support the hypothesis that

DNA-bound NHR family members adopt different conformations to form mutually stabilizing interactions with its partner. It is important to remember that a single bp addition or subtraction imposes a 3.4 angstrom space change and nearly 36 degrees of rotation between the two half complexes.

Because crystal structures are static views of a molecule, pathways of allosteric communication within NHR receptors remain to be understood, i.e. how does ligand-binding affect the rest of the protein. With genomic information becoming readily available, comparative sequence analysis coupled with structure studies were used to address this question. Comparative sequence analysis functions on the premise that aspects of design important for protein function are preserved through evolutionary time. The sequences for 250 NHR ligand binding domains were compared, and 27 residues were identified that link the heterodimerization interface, the ligand binding domain, and two regions important for transmission of signal (Shulman et al., 2004). Mutations of predicted residues disrupted the allosteric network and converted functional heterodimers into non-functional ones. Meanwhile, mutations of residues outside of the network that were spatially close did not cause such a dramatic effect. Another study took a different approach and addressed ligand selectivity between two subtypes of estrogen receptors (Nettles et al., 2004). A series of chimeric estrogen receptors were created by protein domain swaps which led to the identification of regions in estrogen receptors important for ligand selectivity. Structurally-guided point mutation receptor proteins were used to refine the analysis and identified residues within and outside the ligand binding domain to be important for ligand selectivity. Both papers highlight that ligand recognition depends on a network of spatially distributed structural features.

Others have studied how DNA sequence spacing may act as an allosteric regulator for NHR function. This has been best studied with the glucocorticoid receptor (GR), where a connection has been identified between DNA binding sequence, protein conformation, protein dimerization, and transcriptional activity (see review Lefstin and Yamamoto, 1998). The DNA-binding-domain of GR can bind its cognate site in a variety of conformations depending on the sequence of the site, serving to present different interfaces for the recruitment of specific co-factors and to help determine whether the bound factor will activate or repress transcription. Recent work has identified which GR surfaces are required in a gene-specific manner (Rogatsky et al., 2002; Rogatsky et al., 2003). Similar findings have also been reported for the POU family of proteins (see review Phillips and Luisi, 2000). For example, Pit-1, a pituitary-specific POU domain factor, activates growth hormone gene expression in one cell type but restricts expression in another cell type (Scully et al., 2000). This difference is mediated by a two base pair spacing difference in their cognate DNA binding sites that presumably affects the conformation of the bound dimer.

The NHR family has been an excellent system to understand transcription regulation. Where marked DNA-binding rules for other prominent higher eukaryotic transcription factor families have not been identified and may not exist (for example, C/EBP, E2F, and AP-1), the rules for NHR have readily been identified and researchers have used this information to study how receptor-receptor interactions work, what interaction surfaces on the receptors are important for protein-protein interactions, and how changes in protein conformation may result. Lessons learned for NHR will likely be shared with other transcription factor families.

Dorsal

In *Drosophila* development, dorsal-ventral patterning is dependent on a sequence-specific transcription factor Dorsal, a member of the Rel family of transcription factors (Steward, 1987). There are three other Rel family members in *Drosophila*: Dif, Cactus and Relish. While Dorsal is absolutely required for embryogenesis, Dif, Cactus, and Relish have been shown to be dispensible. And while Dorsal is not needed for *Drosophila* immunity, Dif, Cactus, and Relish have been shown to be indispensable. Therefore, Rel family members in *Drosophila* may also have specific and non-redundant functions.

In early cellularizing *Drosophila* embryos, Dorsal protein is distributed in a nuclear gradient with the highest concentrations in ventral regions and lowest concentrations in the lateral and dorsal regions (for a comprehensive review (Stathopoulos and Levine, 2002). This gradient is responsible for the development of mesoderm, neurogenic ectoderm, dorsal epidermis, and amnioserosa. Past studies have demonstrated that Dorsal regulates a network of roughly 25 genes by regulating the amount of Dorsal protein available for DNA binding. Some Dorsal-responsive genes are activated by high levels of dorsal protein (*twist* and *snail*), others are activated by intermediate levels (*sim* and *rhomboid*), and finally some respond only to low levels of dorsal protein (*sog*).

twist is one of the earliest genes activated by the Dorsal gradient and appears in the 12/13 cleavage cycle . It is essential for mesoderm formation and expressed in the 12-14 nuclei located in the ventral-most region of the embryo (Thisse et al., 1991). Analysis of the *twist* promoter identified an enhancer 180 bp upstream from the transcription start site that contained two suboptimal Dorsal

binding sites (Jiang et al., 1991; Thisse et al., 1991). A transgenic fly carrying a promoter-reporter construct consisting of the *twist* enhancer driving *lacZ* expressed *lacZ* in the predicted regions. Mutagenesis experiments changed two base pairs within the Dorsal binding site sequences and converted them into optimal Dorsal binding sites (Jiang and Levine, 1993). A fly carrying this transgenic gene expressed *lacZ* in a much broader range compared with the endogenous gene, with 18-20 nuclei expressing *lacZ*. Thus, the amount of promoter occupancy determined by the Dorsal binding site sequence was integral in establishing different limits of gene expression.

snail is another Dorsal target gene that responds to high levels of Dorsal protein. Snail functions as a repressor in the ventral regions of early embryos, although it is expressed in a more limited area than Twist. Snail's sharp lateral limits of expression are believed to help establish the boundaries between mesoderm and neurogenic ectoderm. Analysis of *snail*'s promoter region revealed binding sites for both Dorsal and Twist (Ip et al., 1992b). The Dorsal binding sites contain suboptimal binding sequences and were located 1-2 kb from the transcription start site, meanwhile the Twist binding sites were located 250 bp from the transcription start site. The sharp pattern of Snail expression may be a direct result of the combinatorial nature of the Dorsal and Twist gradients. Experiments mutating the Twist binding sites in the *snail* promoter resulted in inefficient reporter expression; meanwhile experiments mutating the Dorsal binding sites in the *snail* promoter resulted in no reporter expression (Ip et al., 1992b; Szymanski and Levine, 1995). When Twist binding sites were engineered into the *twist* promoter, reporter expression mimicked Snail expression patterns (Jiang and Levine, 1993). In summary, the Dorsal gradient induces mesoderm development by inducing two genes (more are possible), *twist* and *snail*. *twist* is directly regulated by Dorsal and expressed in the ventral-most region of the embryo. The *snail*

promoter contains binding sites for Dorsal and Twist, and its expression is controlled by the amount of available Dorsal protein and synergy between Dorsal and Twist.

The slope of the Dorsal protein gradient becomes much steeper in the lateral regions of the embryo and allows for as many as five different gradients to be established. Most Dorsal-responsive genes within the lateral regions contain high-affinity Dorsal binding sites. Gene expression patterns are instead largely determined by the location of the Dorsal binding sites and the presence of adjacent transcription factor binding sites. The first major patterning threshold within the neurogenic ectoderm specifies a specialized cell layer called the mesectoderm, a layer of cells that eventually forms specialized neurons and gial cells associated with the ventral nerve cord. These cells depend on Sim expression. Snail blocks the expression of Sim in ventral areas (Kasai et al., 1992). Similar to the *snail* promoter, the *sim* promoter contains binding sites for Dorsal and Twist. Unlike the *snail* promoter, the Dorsal binding sites in the *sim* promoter contain optimal binding sequences and are located much closer to the Twist binding sites as well as the transcription start site (Stathopoulos and Levine, 2002). These three factors may make *sim* a more efficient Dorsal-responsive promoter and explain why Sim expression can occur more laterally than Snail. A recent study has shown that Notch signaling may also impact *sim* expression (Cowden and Levine, 2002). A second patterning region within the neurogenic ectoderm is dependent on Rhomboid expression. Rhomboid expression is controlled by a 300 bp enhancer located 1.7 kb away from the transcription start site (Ip et al., 1992a). Snail blocks the expression of Rhomboid in ventral regions. Analogous to *snail* and *sim*, the *rhomboid* enhancer contains Dorsal and Twist binding sites, but the enhancer also contains binding sites for E-box proteins, another family of bHLH transcription factors. Two E-box proteins, Daughterless and Scute, are ubiquitously expressed

throughout the developing *Drosophila* embryo and have been shown to interact directly with Dorsal (Gonzalez-Crespo and Levine, 1993; Jiang and Levine, 1993). Interactions between Twist-E-box Proteins-Dorsal may further help stabilize Dorsal binding and allow lower levels of Dorsal to function on the *rhomboid* promoter. When the E-Box binding sites in the *rhomboid* promoter were replaced with Twist binding sites, protein expression was narrowed in comparison to wild type (Gray and Levine, 1996). Taken together, interactions of Dorsal with other transcription factors helped establish different thresholds of gene expression within the neurogenic ectoderm.

Finally, the last class of Dorsal responsive genes represented by *sog* (short gastrulation) was identified using a computational method that searched for clusters of Dorsal-binding sites in the *Drosophila* genome (Markstein et al., 2002). The computational method searched for Dorsal binding sites taken from the *zen*, a Dorsal-responsive protein important for ectoderm development (Doyle et al., 1989). Zen expression is normally activated by several ubiquitously expressed transcription factors, but Dorsal represses its expression in the ventral and lateral regions. *zen*'s promoter contains an enhancer region located 1 kb from the transcription start site that contains 3 evenly-spaced Dorsal binding sites and 3 binding sites for Cut and Dead Ringer (cut and dri respectively). Studies showed that Dorsal-Cut-Dri complexes recruit a corepressor complex Groucho, to the promoter to repress gene transcription (Dubnicoff et al., 1997; Valentine et al., 1998). This suggests that interactions between Dorsal-Cut-Dri are integral for *zen* repression. Studies of loss-of-function mutants in Cut or Dri abrogated *zen* repression. Mutagenesis experiments that removed the binding sites for either transcription factor dramatically broadened gene expression patterns, such that protein expression could occur in the ventral-most nuclei. Mutagenesis experiments that replaced Dorsal binding sites in the *zen* promoter with suboptimal

Dorsal binding site sequences taken from *twist* also saw no minimal change in expression patterns (Jiang et al., 1992). This suggested that protein-protein interactions between Dorsal and Cut-Dri were strong enough to overcome any Dorsal affinity barriers. Finally, mutagenesis experiments that replaced the suboptimal Dorsal binding sites in the *twist* promoter with sequences from the *zen* promoter saw a broader range of expression (Jiang and Levine, 1993). This is another example where protein-protein interactions determine expression patterns within the Dorsal gradient. While Dorsal represses *zen* expression, it activates *sog* expression (Markstein et al., 2002). This implies that the Dorsal binding site sequences themselves do not determine whether a gene is activated or repressed. Instead, it is the context where the Dorsal binding sites appear in a promoter (i.e. other transcription factor binding sites or distance away from the transcription start site). The transcription factor complex that Dorsal interacts with on the *sog* promoter still remains to be identified.

All this work has elegantly uncovered how a single transcription factor can coordinate at least five different gene expression programs. High levels of Dorsal activate *twist* and *snail* through suboptimal Dorsal binding sites. Intermediate levels of Dorsal are able to activate *sim*, because the *sim* promoter contains optimal Dorsal and Twist binding sites located much closer to the transcription start site that increases the efficiency of transcription factor binding. *rhomboid* is similarly structured, except a third transcription factor partner is utilized with Dorsal and Twist to establish a slightly different pattern of expression. Finally, low levels of Dorsal regulate *zen* and *sog* through multiple high-affinity Dorsal binding sites and an exceptionally strong interaction partner. It is truly remarkable that a single transcription factor can coordinate so much, and Dorsal highlights the power of coordinate gene expression programs.

In addition to *sog*, the computation approach identified two other functionally-confirmed, novel Dorsal-responsive target genes. The weakness from this first study was the high rate of false positives (77%), but a more recent study that identified neurogenic ectoderm Dorsal-responsive genes had a lower rate (29%) (Markstein et al., 2004). The new study compared more genes within a particular expression class to search for more subtle sequence conservations. In addition to using the *in silico* approach to identifying new Dorsal enhancers. *Drosophila* embryos constitutively expressing high, intermediate, and low levels of Dorsal protein were compared by microarray analysis to identify even more novel Dorsal binding proteins (Stathopoulos et al., 2002). A recent study used information from these genomic studies to tackle functional questions in Rel-mediated *Drosophila* immunity. In depth analysis of the *cecropinA1* promoter, a gene important in *Drosophila* immunity demonstrated that Rel and GATA binding sites were required for gene activation. The same Rel-GATA coupling was seen for other members of the *cecropin* family. Previous microarray studies had identified essential genes involved in response to immunogenic challenge, therefore the authors addressed whether a Rel-GATA linkage was indicative for immunity related genes (De Gregorio et al., 2001). When the top 50 most induced genes from that study were probed for consensus binding sites for Rel and GATA family members, nearly half of the genes showed close linkage between the two sites (Senger et al., 2004). Moreover, a Rel-GATA linkage was also required for expression in larval fat bodies. Promoters that contained isolated Rel or GATA binding sites did not express in larval fat bodies. To demonstrate the synergy between Rel and GATA, a transgenic fly was created that carried a GATA binding site engineered into a normally Rel-responsive, GATA-independent gene. The transgenic gene expressed in larval fat bodies. As seen with *twist*, a few base pair changes within

the DNA sequence altered gene expression patterns. Unlike embryogenesis where Dorsal is the single Rel family member responsible, it still remains to be seen which Rel or GATA family members are required for the Rel-GATA synergy. While a GATA-Rel synergy sounds too simple to explain immunity and larval fat body expression, there are still subtle differences that provide addition regulation. Orientation, spacing between the two sites, and placement relative to the transcription start site were all factors that disrupted specificity and imposed certain restraints.

Specific Genes

The virus inducible interferon-ß promoter has become the standard example of how combinatorial interactions between transcription factors can give rise to a highly specific gene expression program. The IFNß gene is normally silent but is highly induced during viral infection. This behavior is due to an enhancer located in the -110 to -45 region relative to the transcription start site. Three positive regulatory domains have been identified and are recognized by different transcription factor families: NF-κB, IRF, and ATF-2/c-Jun heterodimer. IFNß transcription is induced when each transcription factor is recruited to the promoter in a step-wise fashion. None of the regulatory domains can function alone. Thus, the interferon-ß enhancer acts as an endpoint for many regulatory inputs, which filters and processes information in a way that permits response only to viral infection. Duplicated copies of a particular transcription factor binding site may act as a virus inducible enhancer, but these synthetic enhancers are often less inducible and respond to inducers other than virus infection. In other words, the enhancer is not only a collection of transcription factors binding to their respective sites; the final response comes from synergistic interactions between each of them to form an "enhanceosome." Similar to the SV-40 enhancer, the

IFNß enhancer does not need to be in close proximity to the core promoter. When placed thousands of bases away, the enhancer still functions.

The discovery of HMG I(Y) was an important step to elucidating the mechanisms that controlled the IFNß enhancer. HMG I(Y) is not a typical transcription factor in that it does not directly recruit other proteins. Instead, it unbends and relaxes DNA, lowering the free energy required for activator binding and orchestrating a series of activator-activator interactions. As a result, relatively weak interactions between proteins are strengthened, and the final result is a remarkably stable nucleoprotein complex.

The regulation of the IL-2 gene has also been studied extensively. IL-2 plays a key role in the mammlian immune system by driving the proliferation of B, T, and NK cells. Its expression is T-cell specific and very tightly regulated, such that no basal activity can be detected in resting T cells and a strong induction occurs after T cell activation. How is IL-2 so tightly regulated? IL-2's proximal promoter has been well defined with numerous transcription activation elements including NF-κB, NF-AT, Oct-1, and AP-1 (see review Rothenberg and Ward, 1996). More recent studies have shown that chromatin may play a significant role in IL-2 gene activation. In resting T cells, a nucleosome is positioned over the proximal promoter that prevents the binding of most transcription factors to the promoter (Attema et al., 2002). When T cells are activated, the nucleosome gets remodeled allowing transcription factors to bind their respective sites and activate gene expression. This remodeling is dependent on cRel specifically (Rao et al., 2003). How cRel provides this nonredundant function remains to be understood. However, this is another example

where cRel is specifically required for gene activation; IL-12 production in macrophages was also shown to be cRel dependent as well, although the mechanism is also unclear (Sanjabi et al., 2000).

Chromatin

Chromatin remodeling complexes have been shown to be important in integrating complex promoter regulatory sequences. Further, the expansion of chromatin remodeling complexes and basal transcription machinery over evolutionary time has mimicked the expansion of sequence-specific transcription factors.

While transcription specificity within the NF-κB family has not been significantly addressed in a thorough manner, studies have shown that chromatin remodeling plays an important role in NF-κB dependent gene regulation (Saccani et al., 2001). NF-κB responsive genes have different temporal patterns of gene activation in response to TNF stimulation. By monitoring histone acetylation and p65 binding statuses, NF-κB responsive genes that induced early had constitutive H3 acetylation correlated with p65 binding at early time points. On the other hand, NF-κB responsive genes that induced late had inducible H3 acetylation and p65 did not bind until H3 acetylation had occurred. While this study and cRel's role in IL-2 remodeling only showed correlations between NF-κB and chromatin remodeling, it will be very interesting to understand how and if NF-κB interacts directly with chromatin remodeling complexes.

References

Attema, J. L., Reeves, R., Murray, V., Levichkin, I., Temple, M. D., Tremethick, D. J., and Shannon, M. F. (2002). The human IL-2 gene promoter can assemble a positioned nucleosome that becomes remodeled upon T cell activation. *J. Immunol. 169*, 2466-2476.

Berg, O. G., Winter, R. B., and von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry 20*, 6929-6948.

Cowden, J., and Levine, M. (2002). The Snail repressor positions Notch signaling in the Drosophila embryo. *Development 129*, 1785-1793.

De Gregorio, E., Spellman, P. T., Rubin, G. M., and Lemaitre, B. (2001). Genome-wide analysis of the Drosophila immune response by using oligonucleotide microarrays. *Proc. Natl. Acad. Sci. USA 98*, 12590-12595.

Dodd, I. B., Perkins, A. J., Tsemitsidis, D., and Egan, J. B. (2001). Octamerization of lambda CI repressor is needed for effective repression of P(RM) and efficient switching from lysogeny. *Genes Dev. 15*, 3013-3022.

Dodd, I. B., Shearwin, K. E., Perkins, A. J., Burr, T., Hochschild, A., and Egan, J. B. (2004). Cooperativity in long-range gene regulation by the lambda CI repressor. *Genes Dev. 18*, 344-354.

Doyle, H. J., Kraut, R., and Levine, M. (1989). Spatial regulation of zerknullt: a dorsal-ventral patterning gene in Drosophila. *Genes Dev. 3*, 1518-1533.

Dubnicoff, T., Valentine, S. A., Chen, G., Shi, T., Lengyel, J. A., Paroush, Z., and Courey, A. J. (1997). Conversion of dorsal from an activator to a repressor by the global corepressor Groucho. *Genes Dev. 11*, 2952-2957.

Gonzalez-Crespo, S., and Levine, M. (1993). Interactions between dorsal and helix-loop-helix proteins initiate the differentiation of the embryonic mesoderm and neuroectoderm in Drosophila. *Genes Dev. 7*, 1703-1713.

Gray, S., and Levine, M. (1996). Short-range transcriptional repressors mediate both quenching and direct repression within complex loci in Drosophila. *Genes Dev. 10*, 700-710.

Ip, Y. T., Park, R. E., Kosman, D., Bier, E., and Levine, M. (1992a). The dorsal gradient morphogen regulates stripes of rhomboid expression in the presumptive neuroectoderm of the Drosophila embryo. *Genes Dev. 6*, 1728-1739.

Ip, Y. T., Park, R. E., Kosman, D., Yazdanbakhsh, K., and Levine, M. (1992b). dorsal-twist interactions establish snail expression in the presumptive mesoderm of the Drosophila embryo. *Genes Dev. 6*, 1518-1530.

Jacob, F. (1972). Genetics of the Bacterial Cell. In Nobel Lectures, Physiology or Medicine 1963-1970 (Amsterdam, Elsevier Publishing Company).

Jiang, J., Kosman, D., Ip, Y. T., and Levine, M. (1991). The dorsal morphogen gradient regulates the mesoderm determinant twist in early Drosophila embryos. *Genes Dev. 5*, 1881-1891.

Jiang, J., and Levine, M. (1993). Binding affinities and cooperative interactions with bHLH activators delimit threshold responses to the dorsal gradient morphogen. *Cell 72*, 741-752.

Jiang, J., Rushlow, C. A., Zhou, Q., Small, S., and Levine, M. (1992). Individual dorsal morphogen binding sites mediate activation and repression in the Drosophila embryo. *Embo J. 11*, 3147-3154.

Kalodimos, C. G., Biris, N., Bonvin, A. M., Levandoski, M. M., Guennuegues, M., Boelens, R., and Kaptein, R. (2004). Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. *Science 305*, 386-389.

Kasai, Y., Nambu, J. R., Lieberman, P. M., and Crews, S. T. (1992). Dorsal-ventral patterning in Drosophila: DNA binding of snail protein to the single-minded gene. *Proc. Natl. Acad. Sci. USA 89*, 3414-3418.

Kurokawa, R., Soderstrom, M., Horlein, A., Halachmi, S., Brown, M., Rosenfeld, M. G., and Glass, C. K. (1995). Polarity-specific activities of retinoic acid receptors determined by a co-repressor. *Nature 377*, 451-454.

Lefstin, J. A., and Yamamoto, K. R. (1998). Allosteric effects of DNA on transcriptional regulators. *Nature 392*, 885-888.

Leighton, P. A., Saam, J. R., Ingram, R. S., Stewart, C. L., and Tilghman, S. M. (1995). An enhancer deletion affects both H19 and Igf2 expression. *Genes Dev. 9*, 2079-2089.

Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. *Nature 424*, 147-151.

Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G., and Lu, P. (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science 271*, 1247-1254.

Mader, S., Chen, J. Y., Chen, Z., White, J., Chambon, P., and Gronemeyer, H. (1993). The patterns of binding of RAR, RXR and TR homo- and heterodimers to direct repeats are dictated by the binding specificites of the DNA binding domains. *Embo J. 12*, 5029-5041.

Markstein, M., Markstein, P., Markstein, V., and Levine, M. S. (2002). Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the Drosophila embryo. *Proc. Natl. Acad. Sci. USA 99*, 763-768.

Markstein, M., Zinzen, R., Markstein, P., Yee, K. P., Erives, A., Stathopoulos, A., and Levine, M. (2004). A regulatory code for neurogenic gene expression in the Drosophila embryo. *Development 131*, 2387-2394.

Monod, J. (1972). From Enzymatic Adaption to Allosteric Transitions. In Nobel Lectures, Physiology or Medicine 1963-1970 (Amsterdam, Elsevier Publishing Company).

Nettles, K. W., Sun, J., Radek, J. T., Sheng, S., Rodriguez, A. L., Katzenellenbogen, J. A., Katzenellenbogen, B. S., and Greene, G. L. (2004). Allosteric control of ligand selectivity between estrogen receptors alpha and beta: implications for other nuclear receptors. Mol *Cell 13*, 317-327.

Perlmann, T., Rangarajan, P. N., Umesono, K., and Evans, R. M. (1993). Determinants for selective RAR and TR recognition of direct repeat HREs. *Genes Dev. 7*, 1411-1422.

Phillips, K., and Luisi, B. (2000). The virtuoso of versatility: POU proteins that flex to fit. *J. Mol. Biol. 302*, 1023-1039.

Ptashne, M. (1986). A genetic switch : gene control and phage lambda (Cambridge, Mass.
Palo Alto, Calif., Cell Press ; Blackwell Scientific Publications).

Ptashne, M., and Gann, A. (2002). Genes & signals (Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press).

Rao, S., Gerondakis, S., Woltring, D., and Shannon, M. F. (2003). c-Rel is required for chromatin remodeling across the IL-2 gene promoter. *J. Immunol. 170*, 3724-3731.

Rastinejad, F., Perlmann, T., Evans, R. M., and Sigler, P. B. (1995). Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature 375*, 203-211.

Rastinejad, F., Wagner, T., Zhao, Q., and Khorasanizadeh, S. (2000). Structure of the RXR-RAR DNA-binding complex on the retinoic acid response element DR1. *Embo J. 19*, 1045-1054.

Revet, B., von Wilcken-Bergmann, B., Bessert, H., Barker, A., and Muller-Hill, B. (1999). Four dimers of lambda repressor bound to two suitably spaced pairs of lambda operators form octamers and DNA loops over large distances. *Curr. Biol. 9*, 151-154.

Rogatsky, I., Luecke, H. F., Leitman, D. C., and Yamamoto, K. R. (2002). Alternate surfaces of transcriptional coregulator GRIP1 function in different glucocorticoid receptor activation and repression contexts. *Proc. Natl. Acad. Sci. USA 99*, 16701-16706.

Rogatsky, I., Wang, J. C., Derynck, M. K., Nonaka, D. F., Khodabakhsh, D. B., Haqq, C. M., Darimont, B. D., Garabedian, M. J., and Yamamoto, K. R. (2003). Target-specific utilization of transcriptional regulatory surfaces by the glucocorticoid receptor. *Proc. Natl. Acad.Sci. USA 100*, 13845-13850.

Rothenberg, E. V., and Ward, S. B. (1996). A dynamic assembly of diverse transcription factors integrates activation and cell-type information for interleukin 2 gene regulation. *Proc. Natl. Acad. Sci. USA 93*, 9358-9365.

Saccani, S., Pantano, S., and Natoli, G. (2001). Two waves of nuclear factor kappaB recruitment to target promoters. *J. Exp. Med. 193*, 1351-1359.

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. USA 97*, 12705-12710.

Scully, K. M., Jacobson, E. M., Jepsen, K., Lunyak, V., Viadiu, H., Carriere, C., Rose, D. W., Hooshmand, F., Aggarwal, A. K., and Rosenfeld, M. G. (2000). Allosteric effects of Pit-1 DNA sites on long-term repression in cell type specification. *Science 290*, 1127-1131.

Senger, K., Armstrong, G. W., Rowell, W. J., Kwan, J. M., Markstein, M., and Levine, M. (2004). Immunity regulatory DNAs share common organizational features in Drosophila. *Mol. Cell 13*, 19-32.

Shulman, A. I., Larson, C., Mangelsdorf, D. J., and Ranganathan, R. (2004). Structural determinants of allosteric ligand activation in RXR heterodimers. *Cell 116*, 417-429.

Stathopoulos, A., and Levine, M. (2002). Dorsal gradient networks in the Drosophila embryo. *Dev. Biol*. *246*, 57-67.

Stathopoulos, A., Van Drenth, M., Erives, A., Markstein, M., and Levine, M. (2002). Whole-genome analysis of dorsal-ventral patterning in the Drosophila embryo. *Cell 111*, 687-701.

Steward, R. (1987). Dorsal, an embryonic polarity gene in Drosophila, is homologous to the vertebrate proto-oncogene, c-rel. *Science 238*, 692-694.

Szymanski, P., and Levine, M. (1995). Multiple modes of dorsal-bHLH transcriptional synergy in the Drosophila embryo. *Embo J*. *14*, 2229-2238.

Thisse, C., Perrin-Schmitt, F., Stoetzel, C., and Thisse, B. (1991). Sequence-specific transactivation of the Drosophila twist gene by the dorsal gene product. *Cell 65*, 1191-1201.

Umesono, K., Murakami, K. K., Thompson, C. C., and Evans, R. M. (1991). Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell 65*, 1255-1266.

Valentine, S. A., Chen, G., Shandala, T., Fernandez, J., Mische, S., Saint, R., and Courey, A. J. (1998). Dorsal-mediated repression requires the formation of a multiprotein repression complex at the ventral silencer. *Mol. Cell. Biol*. *18*, 6584-6594.

Webber, A. L., Ingram, R. S., Levorse, J. M., and Tilghman, S. M. (1998). Location of enhancers is essential for the imprinting of H19 and Igf2 genes. *Nature 391*, 711-715.

Zechel, C., Shen, X. Q., Chen, J. Y., Chen, Z. P., Chambon, P., and Gronemeyer, H. (1994). The dimerization interfaces formed between the DNA binding domains of RXR, RAR and TR determine the binding specificity and polarity of the full-length receptors to direct repeats. *Embo J. 13*, 1425-1433.

Zhao, Q., Chasse, S. A., Devarakonda, S., Sierk, M. L., Ahvazi, B., and Rastinejad, F. (2000). Structural basis of RXR-DNA interactions. *J. Mol. Biol*. *296*, 509-520.

CHAPTER 3: GENETIC ANALYSIS OF NF-κB/REL PROTEINS DEFINES FUNCTIONAL SPECIFICITIES WITHIN A MAMMALIAN TRANSCRIPTION FACTOR FAMILY

Alexander Hoffmann, Thomas H. Leung, and David Baltimore

Published in 2003 EMBO 22:5530-5539.

Abstract

The NF-κB transcription factors consist of dimeric proteins of the Rel homology family. They activate many promoters containing highly divergent κB-site sequences. We have generated cell lines lacking individual and multiple NF-κB proteins and used them to establish interactions between components of the NF-κB-IκB signaling system. Functional compensation within the family of dimers was evident in knockout cell lines. Analysis of transiently transfected genes gave an impression of promiscuity that was not borne out by analysis of endogenous genes. Using TNFα as an inducer, a panel of endogenous genes showed a wide a range of subunit specificities as well as highly variable kinetics of induction. Comparing the function and subunit specificity of genes with the sequence of the κB DNA-binding site we found little correlation, indicating that NF-κB family member specificity for endogenous promoters is not solely encoded by the κB site sequence itself.

INTRODUCTION

Cell signaling molecules such as transcription factors are encoded in gene families whose members have distinct yet overlapping functions. The fact that increasingly complex genomes contain a greater number of gene family members rather than novel families (Lander et al., 2001) suggests that diversity within a gene family may provide for specificity and versatility in the regulation of cellular responses. Characterizing the specific functions of the members of transcription factor gene families contributes to our understanding of gene regulatory networks and to the development of specific therapeutic strategies. For example, biochemical studies of the nuclear hormone receptor family have shown that the spacing between two tandem cognate DNA half-sites determines which family members bind and function (Rastinejad, 2001). Similarly within the large zinc finger transcription factor family, biochemical and biophysical characterization of DNA-protein interactions is the basis for synthetic transcription factor engineering (Wolfe et al., 2000). However, for many other transcription factor families, such as the C/EBP (Ramji and Foka, 2002), E2F (Trimarchi and Lees, 2002), AP-1 (Chinenov and Kerppola, 2001; Mechta-Grigoriou et al., 2001), and NF-κB families, marked DNA-binding specificity rules have not been identified. Indeed, the raison d'être of multiple members within signal transduction protein families may not be molecular interaction specificities but could be found in their own differential regulation. Considering the immediate regulators of NF-κB activity, the IκB protein family, for example, the difference between family members that is of primary physiological importance appears to be regulation of synthesis, as demonstrated by the fact that IκBα knockout lethality was rescued by placing IκBß under the control of the IκBα promoter (Cheng et al., 1998). A quantitative analysis of the signaling characteristics of IκB family members has supported that conclusion (Hoffmann et al., 2002).

The NF-κB (also called NF-κB/Rel) family of dimeric transcription factors mediates cellular responses to a wide variety of different stimuli by regulating the expression of a large number of genes of highly diverse functions (Pahl, 1999). Mouse knockout studies have revealed this pleiotropic signal transducer to be involved in inflammatory and stress responses, the control of cell cycle, apoptosis, growth and proliferation, cell-cell communication, lymphocyte development and maturation, and neuronal learning (Gerondakis et al., 1999; Meffert et al., 2003). The functional transcription factors consist of pairs taken from the NF-κB protein family (p50, p52, p65, cRel, RelB), proteins that utilize the Rel homology domain (RHD) for DNA-binding and dimerization. Their cognate DNA-binding element, the κB-site, is found in the promoters of NF-κB responsive genes but displays a remarkably loose consensus, often described as $G_{-5}G_{-4}G_{-3}R_{-2}N_{-1}N_0Y_{+1}Y_{+1}C_{+2}C_{+3}$ (Ghosh et al., 1998). The heterogeneity of κB-sites has been thought to confer specificity of regulation mediated by differential affinities of NF-κB isoforms. Purified recombinant RHD protein dimers could indeed be shown to select differential optimal DNA-binding motifs from a pool of random oligonucleotides (Kunsch et al., 1992). Similarly, one study with transfected promoter-reporter constructs showed that multimerized κB-sites from one of three different promoters (Igκ, MHC H2, and IFN-ß) exhibited differential responses when co-transfected with p50 and/or p65 expression plasmids (Fujita et al., 1992).

However, more recent structural studies did not reveal evidence for highly sequence-specific DNA-NF-κB interactions, with base-specific contacts within the p50 homodimer structure involving primarily 5'-$G_{-5}G_{-4}G_{-3}$ (Ghosh et al., 1995; Muller et al., 1995) and contacts within the p65 homodimer structure involving primarily 5'-$G_{-4}G_{-3}$ (Chen et al., 1998b). In fact, these studies

have not revealed sufficient DNA base contacts to rationalize κB-site sequence-specific binding. Though homodimers of p65 or cRel require only a 9bp sequence for binding, this may be contained within certain 10bp consensus-conforming κB sites. X-ray structures of the p50:p65 heterodimer bound to different κB-sites, such as those derived from the immunoglobulin and HIV enhancers $G_{-5}G_{-4}G_{-3}A_{-2}C_{-1}T_0T_{+1}T_{+2}C_{+3}C_{+4}$ (Chen et al., 1998b), the interferon-ß enhancer $G_{-5}G_{-4}G_{-3}A_{-2}A_{-1}A_0T_{+1}T_{+2}C_{+3}C_{+4}$ (Berkowitz et al., 2002; Escalante et al., 2002), and the urokinase plasminogen activator gene (uPA) promoter $G_{-5}G_{-4}G_{-3}A_{-2}A_{-1}A_0G_{+1}T_{+2}A_{+3}C_{+4}$ (Chen-Park et al., 2002) demonstrate that NF-κB binds a variety of sequences. Furthermore, the two RHD domains contained in each dimer constituent relate to each other differently in different structures and lead to different degrees of DNA bending (Chen and Ghosh, 1999), emphasizing the remarkable permissiveness of NF-κB-DNA interactions.

Despite the difficulty in discerning specificity at the level of DNA-protein interactions, mice deficient in a single RHD protein show specific phenotypes. *rela*$^{-/-}$ animals are embryonic lethal with massive hepatocyte apoptosis (Beg et al., 1995), while *crel*$^{-/-}$ mice develop normally but have defects in lymphocyte proliferation (Kontgen et al., 1995). In part, family member-specific functions can be inferred by tissue specific expression. While p65 appears to be ubiquitous, cRel is expressed constitutively primarily in mature monocytic and lymphocytic lineages (Liou et al., 1994), while in others, such as fibroblasts, only following stimulation (Grumont and Gerondakis, 1990). In addition, there are differences in the molecular characteristics of RHD family members: transfection studies have identified activation domains in p65 and cRel, but not in p50 and p52 (Ghosh et al., 1998). DNA-binding assays were used to identify dimerization rules: p50 (and presumably p52) can partner with all RHD family members including themselves, while RelB and

cRel do not appear to engage in homotypic interactions. RelB, furthermore, does not appear to dimerize with p65 or cRel. Little is known about how the results from *in vitro* biochemical characterizations of RHD proteins can serve to explain physiologically relevant functional specificity in gene regulation, because selectivity of *endogenous* κB-site containing promoters remains largely unexplored. In fact, in the case of the IL-12 promoter, results from cell-free or extra-chromosomal templates are shown not to recapitulate endogenous promoter specificity (Sanjabi et al., 2000).

Genetic approaches are suitable to address questions of *in vivo* mechanistic specificity, but must take into account the possibility of interdependent regulation of factors in mammalian signaling networks; resulting compensatory mechanisms among functionally related molecules may lend increased robustness to signaling systems and affect the phenotype of genetic deletions. Here, we have undertaken a genetic analysis of RHD proteins with respect to NF-κB-dependent gene activation by creating a panel of single and double knockout cell lines. We have biochemically characterized NF-κB activation within them, and measured the TNFα responsiveness of a diverse set of NF-κB target genes. Our results reveal that RHD protein family members exhibit differential target gene specificities and demonstrate that a combination of genetic and biochemical analyses can be applied to decode functional specificity rules of gene promoters in mammalian cells.

RESULTS

Cross-regulation and compensation within the NF-κB-IκB regulatory system

The dominant κB-binding activity induced by TNFα in fibroblasts is composed of p50 and p65, the mature gene products of the *nfκb1* and *rela* genes. Microarray studies with *nfκb1*$^{-/-}$*rela*$^{-/-}$ doubly deficient cells have confirmed their key role in activating most TNFα-responsive genes; fibroblasts lacking both factors show no activation of genes controlled by NF-κB (A.H. and D.B., in preparation). To address the specific roles of p50 and p65 in NF-κB-dependent gene expression, we generated fibroblast cell lines from *nfκb1*$^{-/-}$ (Sha et al., 1995), *rela*$^{-/-}$ (Beg et al., 1995) and *nfκb1*$^{-/-}$*rela*$^{-/-}$ embryos (Horwitz et al., 1997). Western blots confirmed that p50 and p65 proteins were absent in respective knockout cell lines (Fig. #1A), while other RHD proteins remained, though at somewhat altered levels. In particular, p52 levels were enhanced in *nfκb1*$^{-/-}$ cells, cRel levels were increased in *rela*$^{-/-}$ cells, and cRel as well as RelB levels were somewhat reduced in doubly deficient cells. Strikingly, *rela*$^{-/-}$ cells contained markedly reduced levels of IκBα and IκBß proteins, while all three IκB proteins were reduced to almost undetectable levels in doubly deficient cells (Fig. # 1B).

To investigate the result of genetic knockouts within the RHD family on the activated NF-κB transcription factor, we examined the induction of nuclear κB-binding activity in response to TNFα stimulation in wild type and mutant cells lines (Fig. #2). Gelshift and Western blot analyses of the nuclear and cytoplasmic compartments of wild type cells revealed a characteristic biphasic induction of nuclear NF-κB (Fig. #2A) that results from the coordinated degradation and synthesis

of IκB family members (Fig. #2B). Cells lacking p50 showed only a marginal reduction in κB-binding activity (Fig. #2C) with no alteration in the temporal regulation. In p65-deficient cells, however, induction produced two distinct κB-binding activities that did not undergo post-induction repression (Fig. #2E), correlating with defective IκBα protein synthesis (Fig. #2F). Finally, no significant κB-binding activity was induced at any time point within a 6 hour TNFα time course in doubly deficient cells (Fig. #2G). These results confirm that p50 and p65 make up the major constituents of NF-κB in fibroblasts. However, neither p50 nor p65 single knockout cells were deficient in κB-binding activity. In addition, the temporal regulation of nuclear NF-κB is altered in *rela*$^{-/-}$ cells.

Compensation and specificity on non-chromosomal promoter DNA

To characterize the molecular composition of induced κB binding complexes, we employed antibodies specific for each Rel protein in gel mobility supershift analyses of nuclear extracts made after 30 minutes of TNFα stimulation. As expected, nuclear κB-binding activity in wild type cells (Fig. #3A) was sensitive to antibodies specific to p50 and p65; indeed distinct complexes comprising dimers of p65 (upper arrow), p50:p65 (lower arrow) and p50 (data not shown) were identified. The lower complex (indicated by an asterisk) appeared to be non-specific as revealed by competition analysis with wild type and mutant double-stranded oligonucleotides (lanes 2 and 3 in each panel). Not surprisingly, complexes apparent at long exposure in extracts from *nfκb1*$^{-/-}$*rela*$^{-/-}$ double knockout cells (Fig. #3D) were not recognized by any of the NF-κB family antibodies and were not specifically competed by double-stranded oligonucleotides.

Turning to single knockouts, anti-p50 antibodies, as expected, did not recognize κB-binding complexes in extracts derived from *nfκb1*$^{-/-}$ cells (Fig. #3B), nor did anti-p65 antibodies recognize κB-binding complexes derived from *rela*$^{-/-}$ cells (Fig. #3C). Instead, antibodies directed against p52 ablated the *nfκb1*$^{-/-}$ complex of the same mobility as the p50:p65 complex found in wild type cells. Similarly, antibodies directed against cRel supershifted the analogous *rela*$^{-/-}$ κB-binding complex, as well as a much weaker, slower migrating complex that may correspond to a cRel homodimer. These results strikingly revealed molecular compensation within the NF-κB/Rel family; however, not every family member functioned interchangeably with one another. Molecular compensation rules appear to reflect previously noted homology relationships, and the results from molecular characterization studies that inferred structural and functional similarities between p50 and p52, and p65 and cRel proteins (Ghosh et al., 1998).

These results suggest multiple mechanisms of cross-regulation between the RHD and IκB protein family members. One known mechanism is transcriptional feedback: cRel (Grumont et al., 1993), RelB (Bren et al., 2001) and IκBα (Scott et al., 1993) are known NF-κB target genes, which presumably explains observed decreases in protein levels in p50 and p65 doubly deficient cells, as well as the absence of postinduction repression of nuclear NF-κB in *rela*$^{-/-}$ cells. Another cross-regulation mechanism may be based on the protein stability differential between uncomplexed and complexed polypeptides; indeed reduced IκB protein levels in mutant cells may be the result of the loss of NF-κB dimers available for complex formation. Similar reasoning might explain the increased levels of p52 and cRel-containing complexes in *nfκb1*$^{-/-}$ and *rela*$^{-/-}$ cells by respective substitution in dimer formation.

To address the transcriptional activity of p50- or p65-deficient NF-κB complexes, we utilized transient transfections with promoter-reporter constructs driven by multimerized κB-sites. Each construct contained κB-binding site sequences taken from one of three different promoters: Igκ, MHC H2, or interferon-ß. These constructs had previously been shown to have differential responses to co-transfected p50 and p65 expression plasmids (Fujita et al., 1992). We observed significant transcriptional activation of κB-site-containing reporters in wild type fibroblasts following TNFα stimulation (Fig. #4A, red bars). Such activation was dependent on the presence of κB-sites (lower right panel) and was minimal in cells deficient in both p50 and p65 (black bars). However, fibroblasts lacking only one family member, either p50 (green bars) or p65 (blue bars), still showed significant levels of activation, with some reduction observed (about half) in *rela*$^{-/-}$ cells. Surprisingly, the results were similar for all three promoter-constructs. Those κB-sites also revealed little difference in NF-κB complexes when employed in a gelshift assay with nuclear extracts from above-described TNFα-stimulated mutant cells lines. Each probe revealed p50:p65 complexes in wild type cells, p52:p65 complexes in *nfκB1*$^{-/-}$ cells, and p50:cRel complexes in *rela*$^{-/-}$ cells (Fig. #4B, data not shown). However, we note that the H2 κB-probe does not show p65 homodimer complexes in *nfkb1*$^{-/-}$ cells and appears to impart somewhat lower transcriptional activity than the other κB-sites.

The above-described analyses indicate that molecular compensation occurs within the RHD protein family and results in functional compensation on extra-chromosomal templates; the 10bp κB-sites did not reveal much specificity for NF-κB isoforms, though certain sequences may certainly provide a restriction with respect to the entire panel of possible dimers. We next turned to the regulation of *endogenous* genes to determine whether they exhibit similar promiscuity.

However, compensation evident in single knockouts required the inclusion of fibroblast cell lines that are *nfκb1*$^{-/-}$*nfκb2*$^{-/-}$ and *rela*$^{-/-}$*crel*$^{-/-}$ doubly deficient, as well as single knockout controls derived from *nfκb2*$^{-/-}$ and *crel*$^{-/-}$ embryos. For example, to address the function of p65 in wild-type cells, *rela*$^{-/-}$*crel*$^{-/-}$ cells are an informative genotype because the results will not be affected by cRel compensation for the loss of p65. In turn, we can address the question of molecular specificity of RHD proteins in transcriptional activation by determining whether molecular compensation by cRel results in functional compensation on endogenous promoters.

Systematic genetic analysis of a transcription factor family

To undertake such a study, we interbred NF-κB knockout mouse strains and derived 3T3 fibroblast lines that were doubly deficient in p50 and p52 or p65 and cRel. We also derived control lines from nfκb2$^{-/-}$(Caamano et al., 1998) and crel$^{-/-}$ (Kontgen et al., 1995) embryos. Gel mobility shift assays were then used to detect κB-binding complexes in extracts made from each of these cell lines following TNFα stimulation (data not shown); their molecular composition was determined by employing the previously used panel of antibodies. As summarized in Table 1, each mutant cell line showed a characteristic set of TNFα-inducible RHD proteins that gave rise to particular κB-binding activities. Therefore, the systematic panel of mutants allowed us to distinguish between NF-κB-dependent promoters that strictly required a particular RHD protein for activation and those promoters that had a broader requirement for one or the other RHD protein subclass. Specifically, we could distinguish genes that require p50 for TNFα-induced activation, those on which either p50 or p52 must function, and those that do not require either protein for NF-κB-dependent activation. Similarly, those same genes could be further classified according to their requirement for p65 or cRel.

For a quantitative, yet high throughput analysis of endogenous gene expression, we used a multiplex RNAse protection assay that allows independent monitoring of up to 10 different messenger RNA transcripts (Fig. # 5A). Two house keeping genes, GAPDH and L32, allowed for normalization such that consistent expression units can be used to compare different experiments. Multiple data sets from four different wild-type cell lines demonstrated reproducibility, which therefore allows comparison of datasets from different cell lines bearing different mutations. Interestingly, a detailed time course revealed that different genes have widely different kinetics of activation. For example, our panel of genes included several chemokine genes (Fig. # 5B). After TNFα stimulation, MIP-2 mRNA levels increase as early as 15 minutes, reaching a peak at 1 hour, and disappearing by 2 hours. On the hand, RANTES transcripts are not detectable until 2 hours post-stimulation and continue to rise beyond 8 hours. Finally, MCP-1 and IP-10 display an intermediary profile. The dynamic nature of NF-κB activation emphasizes the importance of conducting gene expression assays in detailed time courses.

Differential NF-κB/Rel protein requirements

We focused our attention on 9 genes that represent diverse classes of NF-κB responsive genes, are induced in fibroblasts by TNFα stimulation, and the mRNA levels of which could be monitored reliably by RPA. These genes were: IκBα, aforementioned chemokine genes RANTES, IP10, MCP-1, and MIP-2; the immune regulatory factor LIF; the growth factor M-CSF; transcription factor junB; and apoptosis-inducing TNF receptor family member Fas. Quantitative RPA results for each gene are summarized in three graphs which group related genotypes (Fig. #6). Graphs in the left column demonstrate NF-κB dependence because cells lacking both p50 and p65 have no

NF-κB-mediated activity; those in the center column address dependence on p50 and/or p52 proteins, and those in the right column address dependence on p65 and/or cRel proteins. An eight point time course extending to 8 hours following the onset of TNFα stimulation revealed diverse kinetics in gene activation. In each case, transcriptional induction proved to be protein synthesis independent but cycloheximide did affect the transcript levels of some genes at time points following initial activation (A.H., unpublished results). TNFα activation of all genes proved to be NF-κB-dependent as determined with cells lacking both p50 and p65 (left column). Two genes, M-CSF and junB, showed residual induction in the absence of NF-κB, but the NF-κB contribution to their activation was found to be reproducible and analyzable with respect to NF-κB protein requirement (see below). Interestingly, junB transcription was found to be highly induced in NF-κB knockouts at the time when these cells undergo widespread TNFα-induced apoptosis, indicating that this gene is subject to additional transcriptional regulation mechanisms independent of NF-κB.

Turning to p50 and p52 NF-κB family members, we expected no activation defects in *nfκb2*$^{-/-}$ fibroblasts, because our supershift analysis did not reveal any p52 protein in κB-binding complexes in TNFα stimulated wild type cells. As shown in the center column graphs, that is in fact the case; however, some genes, notably IP-10, appear to be super-induced at later time points in *nfκb2*$^{-/-}$ cells. Surprisingly, TNFα-induced gene expression is largely unaffected by p50 deficiency (green line) as well. In fact, only one gene in our collection, LIF, cannot be activated by TNFα stimulation in *nfκb1*$^{-/-}$ cells. The major reason for such a mild gene expression phenotype appears to be efficient functional compensation by p52, because the activation of several genes, such as

RANTES, IP-10 and M-CSF, is defective in cells lacking both p50 and p52. Thus, NF-κB-dependent genes appear to fall into three categories: those requiring p50 for induction in TNFα-stimulated fibroblasts, those on whose promoters either p50 or p52 can efficiently function, and those whose TNFα induction can proceed without either of these two RHD proteins.

Transcript quantitation for these genes in TNFα-induced cells deficient in cRel, p65, or both is shown in the right column of Fig. #6. Because cRel was not found in κB-binding complexes in wild type fibroblasts, it is not surprising that none of the nine genes are significantly attenuated in $crel^{-/-}$ cells (red lines). In fact, most genes appear to be dramatically dependent on the p65 protein (green lines), although in some cases (e.g. IκBα, LIF, M-CSF) we saw significant transcript levels at the latest time points that are probably the cumulative result of residual transcriptional initiation activity. However, the chemokine MCP-1 and the transcription factor junB did not exhibit such a strict requirement for p65 because cRel could also function on these promoters (compare green and black lines). Importantly, we find that no NF-κB-dependent promoter can be activated in the absence of both p65 and cRel. This observation confirms that these two proteins contain the principal activation functions within κB-binding complexes, while p50 and p52 can be considered binding partners that have essential functions in a subset of gene induction events.

Our panel of knockout cell lines revealed that endogenous genes have remarkably different requirements for RHD proteins when induced by TNFα in fibroblasts. While none could be induced in cells lacking both p50 and p65, or p65 and cRel, the ability to be induced in $nf\kappa B1^{-/-}$, $nf\kappa b1^{-/-}nf\kappa b2^{-/-}$, or $rela^{-/-}$ cells was gene-specific (Table 2, left columns). Therefore, specific RHD protein requirements are determinable for every TNFα induced gene in fibroblasts, as summarized

for the group of nine NF-κB-dependent genes that were part of this study (Table 2, right columns). Within this small sample of genes, the pattern of RHD protein requirement does not appear to correlate with the temporal control of gene expression (indicated in the second column) or the known or presumed function of the gene product; the 6 secreted immune response regulatory proteins represented here (RANTES, MCP-1, IP-10, MIP-2, M-CSF, and LIF) cover diverse temporal kinetics as well as diverse combinations of RHD protein requirements.

DISCUSSION

We have presented genetic evidence that mammalian gene activation requires not only specific transcription factors, but specific members of the RHD transcription factor family for the activation of endogenous NF-κB-responsive genes. However, *in vitro* binding assays with extracts from cells harboring deletions of specific family members reveal DNA-binding characteristics that are sufficiently overlapping that functional specificity of family members cannot be reproduced on naked templates in transient transfections. In fact, in the absence of exogenously expressed family members, transiently transfected promoters appear to be remarkably promiscuous with regard to family member-specific requirement, though some κB-site sequences may be more specialized, restricting access to a subset of NF-κB dimers.

Specificity was revealed when taking a genetic approach to a mechanistic question. The specific RHD protein requirement for the activation of an endogenous promoter implies that a particular subset of all available RHD protein dimers (Table 1) is capable of functioning on the promoter in

question. Thus the RHD protein requirement (Table 2) can be translated into the subset of RHD protein dimers that appear capable of mediating the induction of each gene in our study (Table 3). For example, LIF induction by TNFα is abolished in both *nfκb1*$^{-/-}$ as well as *rela*$^{-/-}$ single knockout cells leading to the conclusion that of all possible RHD protein dimers only the p50:p65 heterodimer is functional on this promoter. In contrast, IP-10 (as well as M-CSF, and RANTES) can be induced by dimers containing either p50 or p52, with p65 being a required partner. A representative of a third group is MIP-2; this chemokine (as well as IκBα and Fas) is induced by p65-containing complexes that may not only be heterodimers with p50 or p52, but can also be p65:p65 homodimers. MCP-1 and junB exhibited the least stringent RHD protein requirement with NF-κB/Rel complexes consisting of either p65 or cRel capable of functioning on these promoters. While cRel protein can partner with p50, p65 can function as a homodimer or as a heterodimer with p50 or p52 to activate transcription of these two genes.

Determinants of NF-κB protein requirements

Because individual NF-κB dimers have been shown to have particular binding sequence preferences *in vitro* (Kunsch et al., 1992), we examined whether RHD protein requirements can be correlated with the sequence of the κB-site. Such a correlation would indicate that DNA-protein interaction affinity, or more precisely K_{on} and/or k_{off} rates, of RHD protein dimers to specific κB site sequences determine which RHD protein dimers are functional on a particular promoter. The promoters for most of the genes in this study have been previously investigated by standard transiently transfected reporter assays that led to the identification of the apparent functional transcription factor binding sites. For LIF, we used human and mouse genomic databases and previous promoter studies to identify a novel κB-binding element. Grouping the NF-κB-dependent

genes in our study according their specific RHD protein requirement in Table 3 failed to reveal a correlation with κB-site sequences. In fact, genes whose TNFα induction can be mediated by the same set of RHD protein dimers, such as RANTES, IP-10, M-CSF, contain different κB-site sequences that are responsible for this induction. Conversely, promoters containing the same κB-site sequence (GGGAATTTCC), such as MIP-2 and MCP-1, exhibit different RHD protein requirements. Furthermore, focusing on genes that can be activated by p65 homodimers (MCP-1, junB, MIP-2, IκBα and Fas) we cannot discern any common features in the relevant κB-site sequences. In each case, the κB element is embedded in a promoter that is bound by many proteins, such as neighboring transcription factors, co-activators, and chromatin components that may interact with NF-κB/Rel dimers bound at the κB-site. The results of our genetic analysis suggest that the ability of RHD protein dimers to function on a particular promoter is not solely determined by DNA-protein interactions at the κB-site. While the regulatory control regions of the genes in this study have previously been characterized, we cannot rule out that NF-κB may function via unrecognized binding sites, whose sequence may not be conserved between mouse and human. Mutations of endogenous sequences via knock-in technology or methodologies that allow faithful expression of transgenic constructs will be needed to confirm and extend our conclusions.

Interactions with DNA is likely to restrict the subset of possible transcription factors that may function at a particular promoter element, and within the family of κB sites it is known that p65 or cRel homodimers bind better to 9bp than 10bp sites. However, our results (based on the highly divergent set of 10bp κB sites) point to the importance of protein-protein interactions within promoter-DNA assemblies in determining a specific family member requirement and generating

transcriptional specificity *in vivo*. Transcriptional synergy (Lin et al., 1990) between adjacent transcription factors and within enhanceosomes (Thanos and Maniatis, 1995b), as well as specific co-activator requirements (Merika et al., 1998), have been observed in non-chromosomal experimental systems and have long been thought to play a role in generating specificity through combinatorial control. Differential specificity of NF-κB protein family members in interactions with contextual transcriptional factors may thus account for family member specific requirement for gene activation. Chromatin has also been shown to control NF-κB/Rel accessibility *in vivo* (Saccani et al., 2001) and may in fact do so in a manner that is specific for a subset of RHD protein dimers. Furthermore, some genes require chromatin reorganization for gene activation (Lomvardas and Thanos, 2002), and this may be dependent on protein-protein interactions specific to a particular family member.

Other groups have taken complementary approaches to studying transcription factor specificity. In particular, Farnham and colleagues used chromatin immunoprecipitation in combination with microarray technology (ChIP-Chip) to examine the role of a single member of the E2F family across a wide spectrum of genes (Weinmann et al., 2002), while Young and colleagues have begun to characterize extensive networks of transcriptional control in budding yeast (Lee et al., 2002). Recently, ChIP experiments have revealed differential NF-κB dimer recruitment to target promoters that may complement the present study (Saccani et al., 2003). While NF-κB occupancy as assayed by ChIP does not indicate the functional activity of the bound dimer without the use of genetic tools, such assays indicate that transcriptional control may involve the dynamic interplay of different dimers to attenuate or sustain transcriptional activity (also A.H. and D.B., manuscript in preparation). Perfecting the ChIP-chip technology to allow for a comprehensive characterization

of occupancy states of regulatory regions will shed light on promoter architecture. Complementing that approach with transcript measurements in a variety of relevant knockouts may illuminate the operative promoter code, and in particular the question of family member specificity.

Are distinct κB-site sequences functionally equivalent? While our genetic analysis has failed to correlate functional requirement with κB site sequence, and NF-κB co-crystal structure studies do not reveal extensive NF-κB dimer-specific DNA-binding contacts (Chen and Ghosh, 1999), this does not rule out that the κB-site sequence plays a functional role. In fact, structural comparisons of the p50:p65 dimer bound to different κB-site sequences have revealed significant differences in conformation (Chen and Ghosh, 1999). The particular conformations of DNA bound NF-κB complexes (probed by protease sensitivity and X-ray crystallography) have indeed between correlated with their ability to transactivate on transfected promoters (Chen-Park et al., 2002; Fujita et al., 1992); however it is unknown whether the conformational state of NF-κB complexes may have similar importance on chromosomal promoters, and whether a particular conformation of an NF-κB dimer may be functional on one promoter but not another. In addition to examining binding affinities, investigating the conformations of NF-κB dimers bound to κB elements and the functional consequence within the context of endogenous regulatory region is likely to bring further understanding to the promoter code.

Multiple protein-DNA complexes may be able to assemble on a given promoter depending on the cell-type and stimulus (Falvo et al., 2000). Because the κB-site sequence is not the sole determinant of the specific RHD dimer requirement, we may find that the specific RHD protein requirement for the activation of a particular gene is also specific to cell type and stimulus. If

combinatorial control by multiple promoter-bound transcription factors generates specific transcription factor requirements, it can form the basis for the inclusion or exclusion of a particular gene in specific gene expression programs in response to multiple activation pathways.

MATERIALS AND METHODS

Cell culture and transfections

Immortalized fibroblast cell lines were generated from E12.5-14.5 embryos according to the 3T3 protocol in 10% bovine calf serum (Aaronson and Todaro, 1968) and maintained in the same manner. Previously published NF-κB luciferase reporter plasmids (Fujita et al., 1992) were transfected together with ß-actin-lacZ control plasmid into subconfluent fibroblasts using Ca-PO_4 or fugene6 (Roche) methodologies. After 24hrs, cells were starved for 24hrs with 0.5% serum containing medium and then stimulated with 10ng/ml TNFα.

DNA-binding assays and Western blot

For whole cell Western blots, cells were lysed within 6 well plates using SDS-PAGE sample buffer, and probed with specific antibodies from Santa-Cruz Biotechnology, Inc. For EMSA, previously described methods and κB-site containing probes were used (Fujita et al., 1992; Hoffmann et al., 2002). Supershift assays were done with antibody cocktails specific to indicated RHD proteins composed of rabbit antibodies available at Santa Cruz Biotechnology. The reactivity and specificity of these is demonstrated in the data shown in Fig. #3A-D.

RNAse protection assays

Total RNA was made from confluent and starved fibroblasts using Tri-Reagent (Molecular Research Center, Inc). RNase Protection Analysis was performed with 5µg RNA using Riboquant probe sets (Pharmingen) according to manufacturer's instructions. Data was quantitated using a Molecular Dynamics Phorphorimager. Following local background subtraction, data was normalized using GAPDH and L32 gene data, which allowed results to be compared across

experiments. Every experiment was performed at least twice, many 3 to 5 times with high reproducibility.

ACKNOWLEDGMENTS

A.H. thanks A. Beg, W.C. Sha, B.H. Horwitz, S. Gerondakis, R. Ryseck, and R. Bravo for knockout mice, L. Anonueovo, B. Kennedy, and S. Pease for their maintenance, S. Tronick for antibodies, and G. Ghosh and T. Huxford for critical reading of the manuscript. T.H.L. is an ARCS Foundation fellow.

References

Aaronson, S.A. and Todaro, G.J. (1968) Development of 3T3-like lines from Balb-c mouse embryo cultures: transformation susceptibility to SV40. *J. Cell. Physiol.*, 72, 141-148.

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

Berkowitz, B., Huang, D.B., Chen-Park, F.E., Sigler, P.B. and Ghosh, G. (2002) The x-ray crystal structure of the NF-kappa B p50.p65 heterodimer bound to the interferon beta -kappa B site. *J. Biol. Chem.*, 277, 24694-24700.

Bren, G.D., Solan, N.J., Miyoshi, H., Pennington, K.N., Pobst, L.J. and Paya, C.V. (2001) Transcription of the RelB gene is regulated by NF-kappaB. *Oncogene*, 20, 7722-7733.

Caamano, J.H., Rizzo, C.A., Durham, S.K., Barton, D.S., Raventos-Suarez, C., Snapper, C.M. and Bravo, R. (1998) Nuclear factor (NF)-kappa B2 (p100/p52) is required for normal splenic microarchitecture and B cell-mediated immune responses. *J. Exp. Med.*, 187, 185-196.

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

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

Chen-Park, F.E., Huang, D.B., Noro, B., Thanos, D. and Ghosh, G. (2002) The kappa B DNA sequence from the HIV long terminal repeat functions as an allosteric regulator of HIV transcription. *J. Biol. Chem.*, 277, 24701-24708.

Cheng, J.D., Ryseck, R.P., Attar, R.M., Dambach, D. and Bravo, R. (1998) Functional redundancy of the nuclear factor kappa B inhibitors I kappa B alpha and I kappa B beta. *J. Exp. Med.*, 188, 1055-1062.

Chinenov, Y. and Kerppola, T.K. (2001) Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. *Oncogene*, 20, 2438-2452.

Escalante, C.R., Shen, L., Thanos, D. and Aggarwal, A.K. (2002) Structure of NF-kappaB p50/p65 heterodimer bound to the PRDII DNA element from the interferon-beta promoter. *Structure (Camb)*, 10, 383-391.

Falvo, J.V., Uglialoro, A.M., Brinkman, B.M., Merika, M., Parekh, B.S., Tsai, E.Y., King, H.C., Morielli, A.D., Peralta, E.G., Maniatis, T., Thanos, D. and Goldfeld, A.E. (2000) Stimulus-specific assembly of enhancer complexes on the tumor necrosis factor alpha gene promoter. *Mol. Cell. Biol.*, 20, 2239-2247.

Fujita, T., Nolan, G.P., Ghosh, S. and Baltimore, D. (1992) Independent modes of transcriptional activation by the p50 and p65 subunits of NF-kappa B. *Genes Dev.*, 6, 775-787.

Gerondakis, S., Grossmann, M., Nakamura, Y., Pohl, T. and Grumont, R. (1999) Genetic approaches in mice to understand Rel/NF-kB and IkB function: trangenics and knockouts. *Oncogene*, 18, 6888-6895.

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.

Ghosh, S., May, M.J. and Kopp, E.B. (1998) NF-kB and Rel proteins: Evolutionarily conserved mediators of immune responses. *Annu. Rev. Immunol.*, 16, 225-260.

Grumont, R.J. and Gerondakis, S. (1990) Murine c-rel transcription is rapidly induced in T-cells and fibroblasts by mitogenic agents and the phorbol ester 12-O-tetradecanoylphorbol-13-acetate. *Cell. Growth Differ.*, 1, 345-350.

Grumont, R.J., Richardson, I.B., Gaff, C. and Gerondakis, S. (1993) rel/NF-kappa B nuclear complexes that bind kB sites in the murine c-rel promoter are required for constitutive c-rel transcription in B-cells. *Cell. Growth Differ.*, 4, 731-743.

Hoffmann, A., Levchenko, A., Scott, M.L. and Baltimore, D. (2002) The IkB-NF-kB signaling module: temporal control and selective gene activation. *Science*, 298, 1241-1245.

Horwitz, B.H., Scott, M.L., Cherry, S.R., Bronson, R.T. and Baltimore, D. (1997) Failure of lymphopoiesis after adoptive transfer of NF-kappaB-deficient fetal liver cells. *Immunity*, 6, 765-772.

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

Kunsch, C., Ruben, S.M. and Rosen, C.A. (1992) Selection of optimal 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.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S.,

Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., Gibbs, R.A., Muzny, D.M., Scherer, S.E., Bouck, J.B., Sodergren, E.J., Worley, K.C., Rives, C.M., Gorrell, J.H., Metzker, M.L., Naylor, S.L., Kucherlapati, R.S., Nelson, D.L., Weinstock, G.M., Sakaki, Y., Fujiyama, A., Hattori, M., Yada, T., Toyoda, A., Itoh, T., Kawagoe, C., Watanabe, H., Totoki, Y., Taylor, T., Weissenbach, J., Heilig, R., Saurin, W., Artiguenave, F., Brottier, P., Bruls, T., Pelletier, E., Robert, C., Wincker, P., Smith, D.R., Doucette-Stamm, L., Rubenfield, M., Weinstock, K., Lee, H.M., Dubois, J., Rosenthal, A., Platzer, M., Nyakatura, G., Taudien, S., Rump, A., Yang, H., Yu, J., Wang, J., Huang, G., Gu, J., Hood, L., Rowen, L., Madan, A., Qin, S., Davis, R.W., Federspiel, N.A., Abola, A.P., Proctor, M.J., Myers, R.M., Schmutz, J., Dickson, M., Grimwood, J., Cox, D.R., Olson, M.V., Kaul, R., Shimizu, N., Kawasaki, K., Minoshima, S., Evans, G.A., Athanasiou, M., Schultz, R., Roe, B.A., Chen, F., Pan, H., Ramser, J., Lehrach, H., Reinhardt, R., McCombie, W.R., de la Bastide, M., Dedhia, N., Blocker, H., Hornischer, K., Nordsiek, G., Agarwala, R., Aravind, L., Bailey, J.A., Bateman, A., Batzoglou, S., Birney, E., Bork, P., Brown, D.G., Burge, C.B., Cerutti, L., Chen, H.C., Church, D., Clamp, M., Copley, R.R., Doerks, T., Eddy, S.R., Eichler, E.E., Furey, T.S., Galagan, J., Gilbert, J.G., Harmon, C., Hayashizaki, Y., Haussler, D., Hermjakob, H., Hokamp, K., Jang, W., Johnson, L.S., Jones, T.A., Kasif, S., Kaspryzk, A., Kennedy, S., Kent, W.J., Kitts, P., Koonin, E.V., Korf, I., Kulp, D., Lancet, D., Lowe, T.M., McLysaght, A., Mikkelsen, T., Moran, J.V., Mulder, N., Pollara, V.J., Ponting, C.P., Schuler, G., Schultz, J., Slater, G., Smit, A.F., Stupka, E., Szustakowski, J., Thierry-Mieg, D., Thierry-Mieg, J., Wagner, L., Wallis, J., Wheeler, R., Williams, A., Wolf, Y.I., Wolfe, K.H., Yang, S.P., Yeh, R.F., Collins, F., Guyer, M.S., Peterson, J., Felsenfeld, A., Wetterstrand, K.A., Patrinos, A., Morgan, M.J., Szustakowki, J., de Jong, P., Catanese, J.J., Osoegawa, K., Shizuya, H., Choi, S. and Chen, Y.J. (2001) Initial sequencing and analysis of the human genome. *Nature*, 409, 860-921.

Lee, T.I., Rinaldi, N.J., Robert, F., Odom, D.T., Bar-Joseph, Z., Gerber, G.K., Hannett, N.M., Harbison, C.T., Thompson, C.M., Simon, I., Zeitlinger, J., Jennings, E.G., Murray, H.L., Gordon, D.B., Ren, B., Wyrick, J.J., Tagne, J.B., Volkert, T.L., Fraenkel, E., Gifford, D.K. and Young, R.A. (2002) Transcriptional regulatory networks in Saccharomyces cerevisiae. *Science*, 298, 799-804.

Lin, Y.S., Carey, M., Ptashne, M. and Green, M.R. (1990) How different eukaryotic transcriptional activators can cooperate promiscuously. *Nature*, 345, 359-361.

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

Lomvardas, S. and Thanos, D. (2002) Modifying gene expression programs by altering core promoter chromatin architecture. *Cell*, 110, 261-271.

Mechta-Grigoriou, F., Gerald, D. and Yaniv, M. (2001) The mammalian Jun proteins: redundancy and specificity. *Oncogene*, 20, 2378-2389.

Meffert, M.K., Chang, J.M., Wiltgen, B.J., Fanselow, M.S. and Baltimore, D. (2003) NF-kB Functions in Synaptic Signaling and Behavior. *Nature Neuroscience*.

Merika, M., Williams, A.J., Chen, G., Collins, T. and Thanos, D. (1998) Recruitment of CBP/p300 by the IFN beta enhanceosome is required for synergistic activation of transcription. *Mol. Cell*, 1, 277-287.

Muller, C.W., Rey, F.A., Sodeoka, M., Verdine, G.L. and Harrison, S.C. (1995) Structure of the NF-kappa B p50 homodimer bound to DNA. *Nature*, 373, 311-317.

Pahl, H.L. (1999) Activators and target genes of Rel/NF-kB transcription factors. *Oncogene*, 18, 6853-6866.

Ramji, D.P. and Foka, P. (2002) CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem. J.*, 365, 561-575.

Rastinejad, F. (2001) Retinoid X receptor and its partners in the nuclear receptor family. *Curr. Opin. Struct. Biol.*, 11, 33-38.

Saccani, S., Pantano, S. and Natoli, G. (2001) Two waves of nuclear factor kappaB recruitment to target promoters. *J. Exp. Med.*, 193, 1351-1359.

Saccani, S., Pantano, S. and Natoli, G. (2003) Modulation of NF-kappaB activity by exchange of dimers. *Mol. Cell.*, 11, 1563-1574.

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. USA*, 97, 12705-12710.

Scott, M.L., Fujita, T., Liou, H.C., Nolan, G.P. and Baltimore, D. (1993) The p65 subunit of NF-kappa B regulates I kappa B by two distinct mechanisms. *Genes Dev.*, 7, 1266-1276.

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.

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

Trimarchi, J.M. and Lees, J.A. (2002) Sibling rivalry in the E2F family. *Nat. Rev. Mol. Cell. Biol.*, 3, 11-20.

Weinmann, A.S., Yan, P.S., Oberley, M.J., Huang, T.H. and Farnham, P.J. (2002) Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev.*, 16, 235-244.

Wolfe, S.A., Nekludova, L. and Pabo, C.O. (2000) DNA recognition by Cys2His2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct.*, 29, 183-212.

FIGURE LEGENDS

Fig. #1 RHD proteins in NF-κB knockout cells.

Western blots of cells derived from wild-type, *nfκB1*$^{-/-}$, *relA*$^{-/-}$ and *nfκB1*$^{-/-}$/*relA*$^{-/-}$ mouse embryos with antibodies directed against p50, p52, p65, cRel, RelB, and actin (A), and IκBε, -ß, -α, and actin (B) as indicated. Specific bands are shown, but additional bands were detectable with some antibodies.

Fig. #2 TNFα induction of κB-binding activity in NF-κB knockout cells.

Electrophoretic mobility shift assay (EMSA) was used to monitor nuclear NF-κB in wild type (A), *nfκB1*$^{-/-}$ (C), *relA*$^{-/-}$ (E) and *nfκB1*$^{-/-}$*relA*$^{-/-}$ (G) cells in the indicated (in minutes and hours) timecourse following the onset of TNFα stimulation. The cytoplasmic portion from wild type (B), *nfκB1*$^{-/-}$ (D), *relA*$^{-/-}$ (F) was probed in Western blots to monitor concurrent degradation and synthesis of IκB proteins, IκBα and IκBß (bottom panels), and IκBε (top panels).

Fig. #3 Molecular composition of κB-binding activity in NF-κB knockout cells.

Immediate early protein-DNA complexes induced by TNFα in wild type (A), *nfκB1*$^{-/-}$ (B), *relA*$^{-/-}$ (C) and *nfκB1*$^{-/-}$*relA*$^{-/-}$ (D) cells and detected by EMSA in Figure 2 are indicated by arrows in lanes 1, while constitutive complexes are indicated by asterisks. They are specific for κB-sites as shown by competition with double stranded wild type and mutant oligonucleotides (lanes 2 and 3) and are probed with antibodies directed against the indicated RHD proteins (lanes 4-9). This results in "supershift" or ablation of the characteristic protein-DNA complex. We conclude that the κB-binding activity in wild type cells consists of p50:p65 heterodimer and p65 homodimer, in *nfκB1*$^{-/-}$

cells p52:p65 heterodimer and p65 homodimer, in *rela*$^{-/-}$ cells p50 homodimer, cRel:p50 heterodimer, and most likely, but not unambiguously cRel homodimer. While experiments shown in panels A, B, C are exposed to film for 6 hrs, panel D shows a 24hr exposure.

Fig. #4 Compensation on different κB-site sequences in extra-chromosomal plasmids.

Reporter plasmids driven by the c-fos core promoter alone or fused to two κB sites derived from the Igκ the MHC H2, and IFNß promoter, as indicated in (A), were assayed in response to TNFα stimulation for the indicated time in wild type (red bars), *nfκB1*$^{-/-}$ (green bars), *relA*$^{-/-}$ (blue bars) and *nfκB1*$^{-/-}$*relA*$^{-/-}$ (black bars). EMSA (B) was used to monitor κB-binding activity in TNFα stimulated NF-κB knockout cells with indicated κB-site containing double-stranded oligo nucleotide probes. Induced, specific bands are indicated by arrows, a prominent constitutive protein-DNA complex is indicated by an asterisk.

Table 1. RHD proteins induced by TNFα stimulation in NF-κB knockout cells.

Fibroblast cell lines derived from embryos of indicated genotypes (column 1) exhibit TNFα inducible κB binding activity that consists of indicated RHD proteins (column 2). These proteins form κB-binding complexes indicated in column 3. All data are based on EMSA with κB-site containing probes and antibodies directed against specific RHD proteins. Complexes marked with an asterisk were detected with only a subset of κB-site containing probes.

Fig. # 5 Quantitative monitoring of gene expression by multiplex RNAse protection assay (RPA).

Multiple double stranded RNA products indicative of mRNA transcripts derived from indicated NF-κB-dependent and house keeping genes following TNFα stimulation of wild type fibroblasts

resolved by electrophoresis (A). Three independent experiments with independently derived wild type 3T3 cell lines were quantitated by phospho-imager and mRNA abundance was graphed for the indicated chemokine genes in arbitrary units relative to housekeeping genes L32 and GAPDH (B).

Fig. #6 Transcriptional induction by TNFα of NF-κB target genes in NF-κB knockout cells.

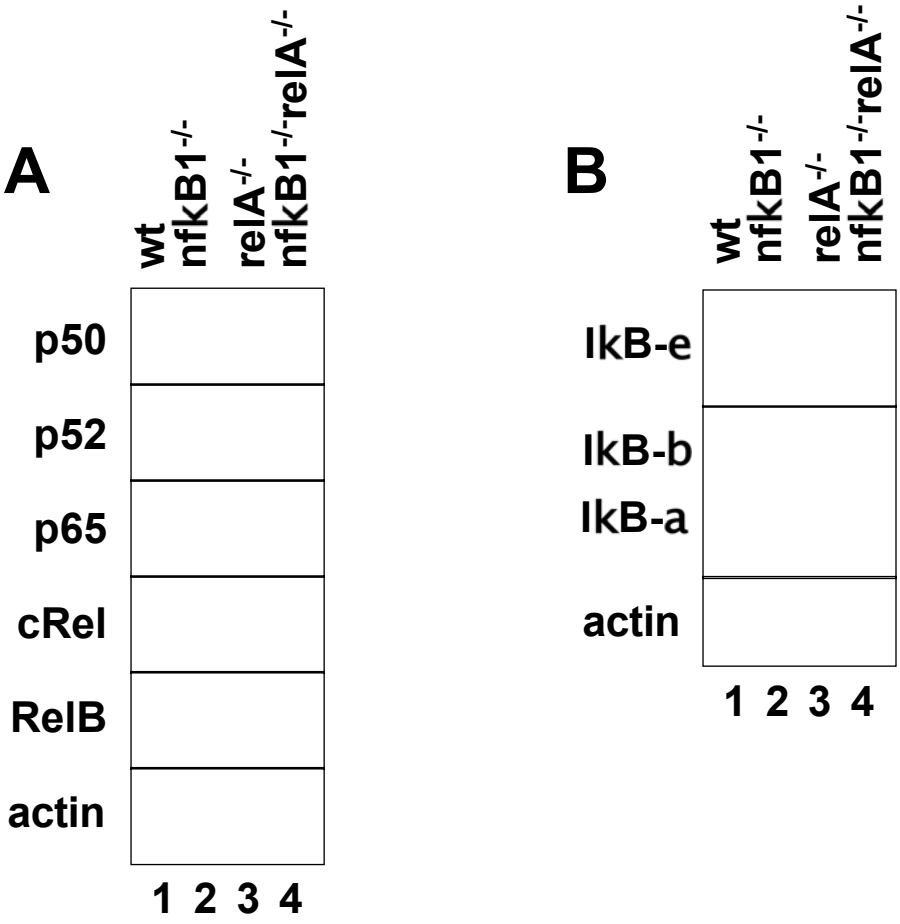
Results from a representative RPA experiment are graphed for the indicated genes in three panels. The left panel shows data from wild type (red line) and *nfκb1*$^{-/-}$*rela*$^{-/-}$ cells (black line), the center panel from *nfκb2*$^{-/-}$ (red), *nfκb1*$^{-/-}$ (green), and *nfκb1*$^{-/-}$*nfκb2*$^{-/-}$ (black), and the right panel from *crel*$^{-/-}$ (red), *rela*$^{-/-}$ (green), and *rela*$^{-/-}$*crel*$^{-/-}$ (black) cells.

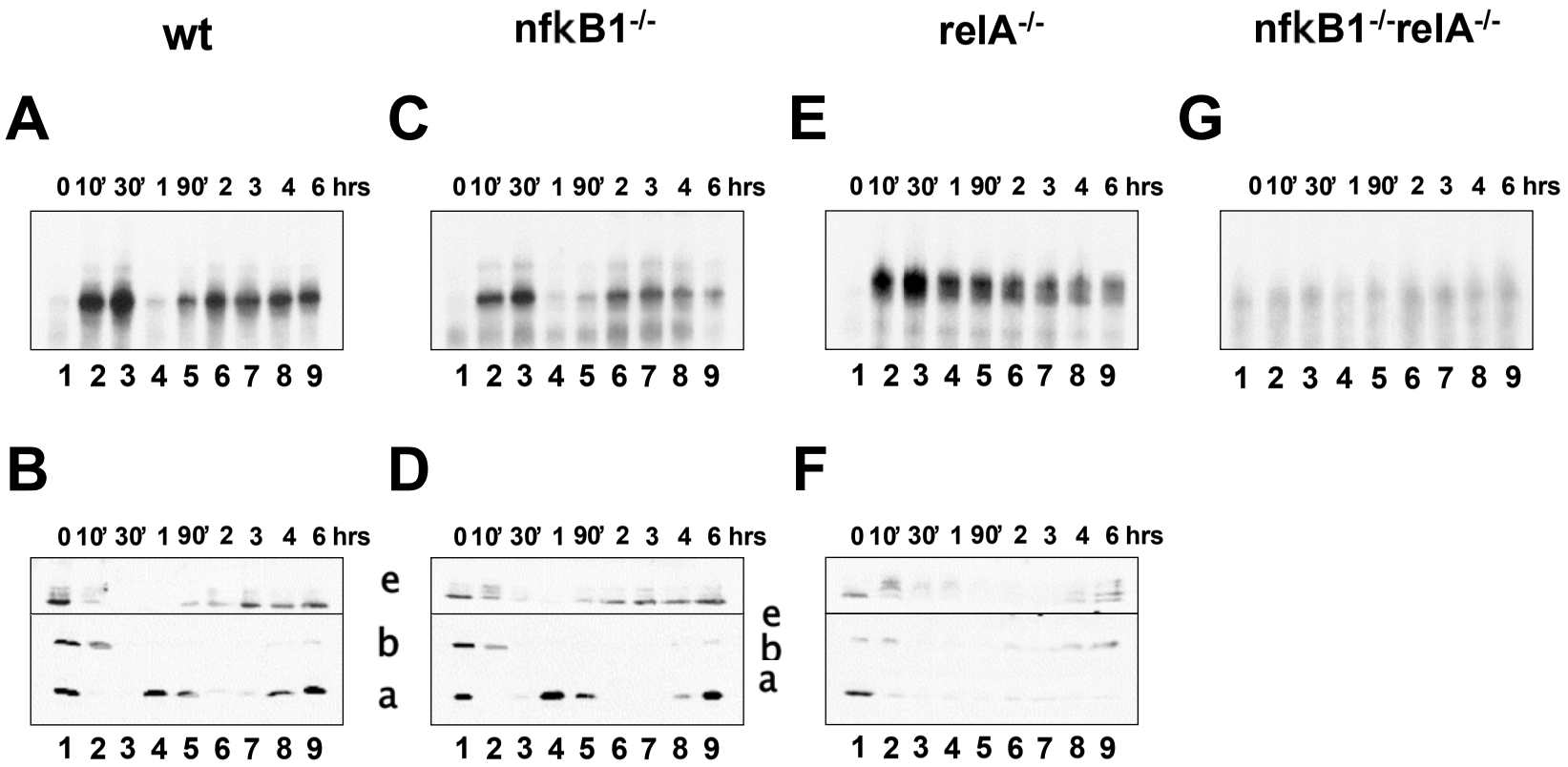
Table 2. Distinct RHD protein requirements for NF-κB-dependent genes.

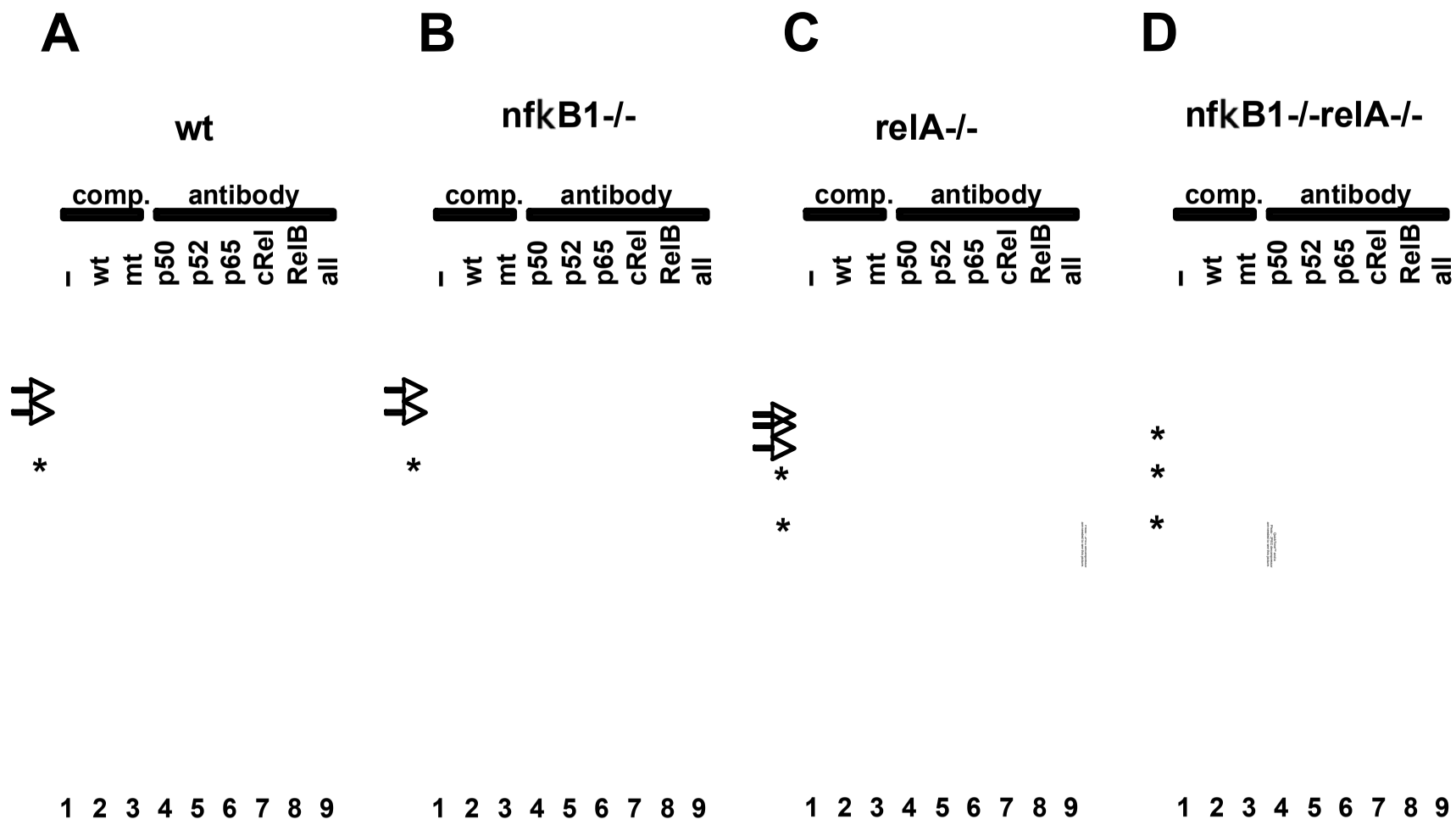
Transcriptional regulation results are summarized for each of the ten genes (column 1) monitored: Temporal regulation (column 2), expression in indicated knockout cells (columns 3-5), and deduced RHD protein requirement (column 6 - 7).

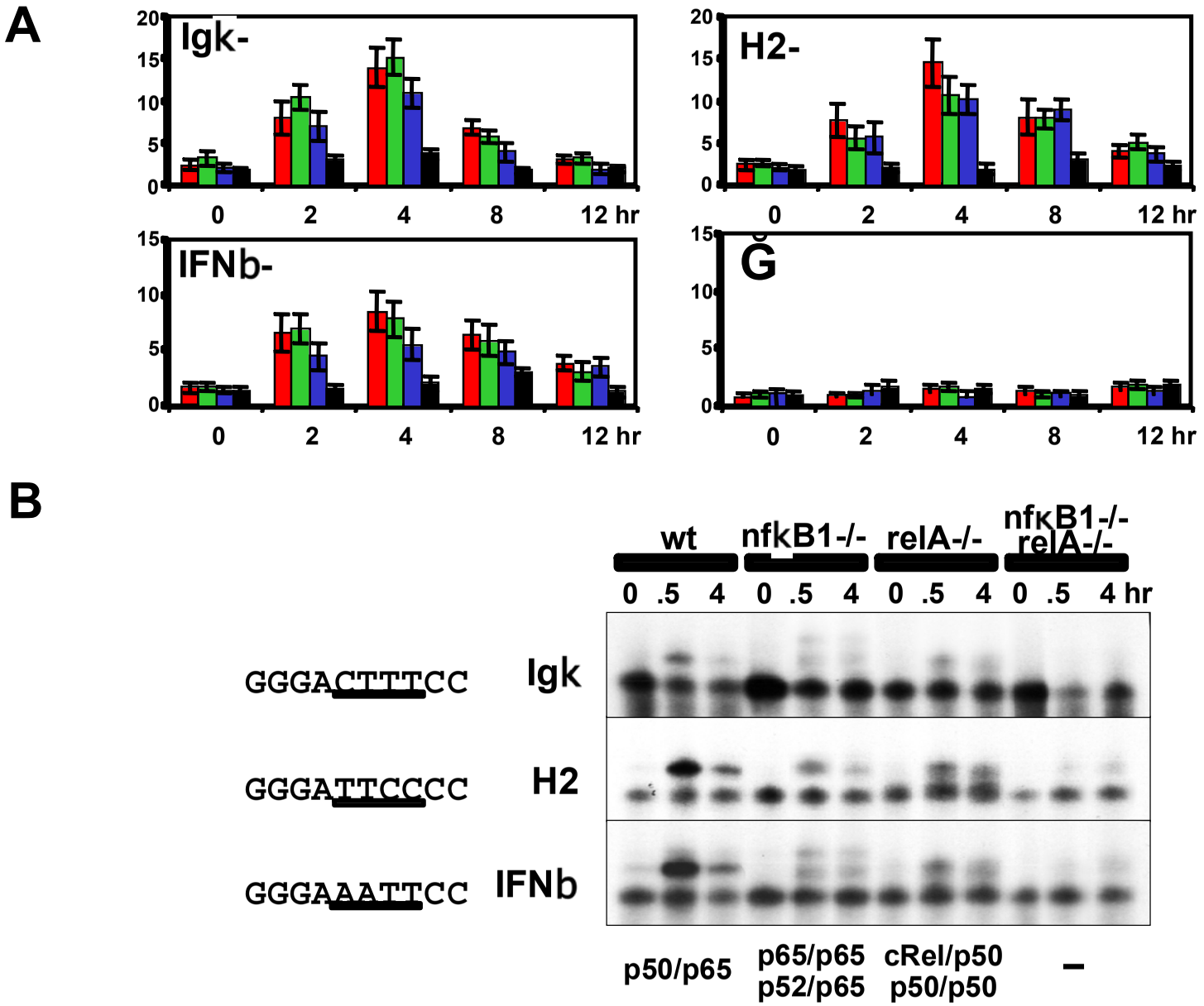
Table 3. κB-site sequences do not correlate with RHD dimer specificity.

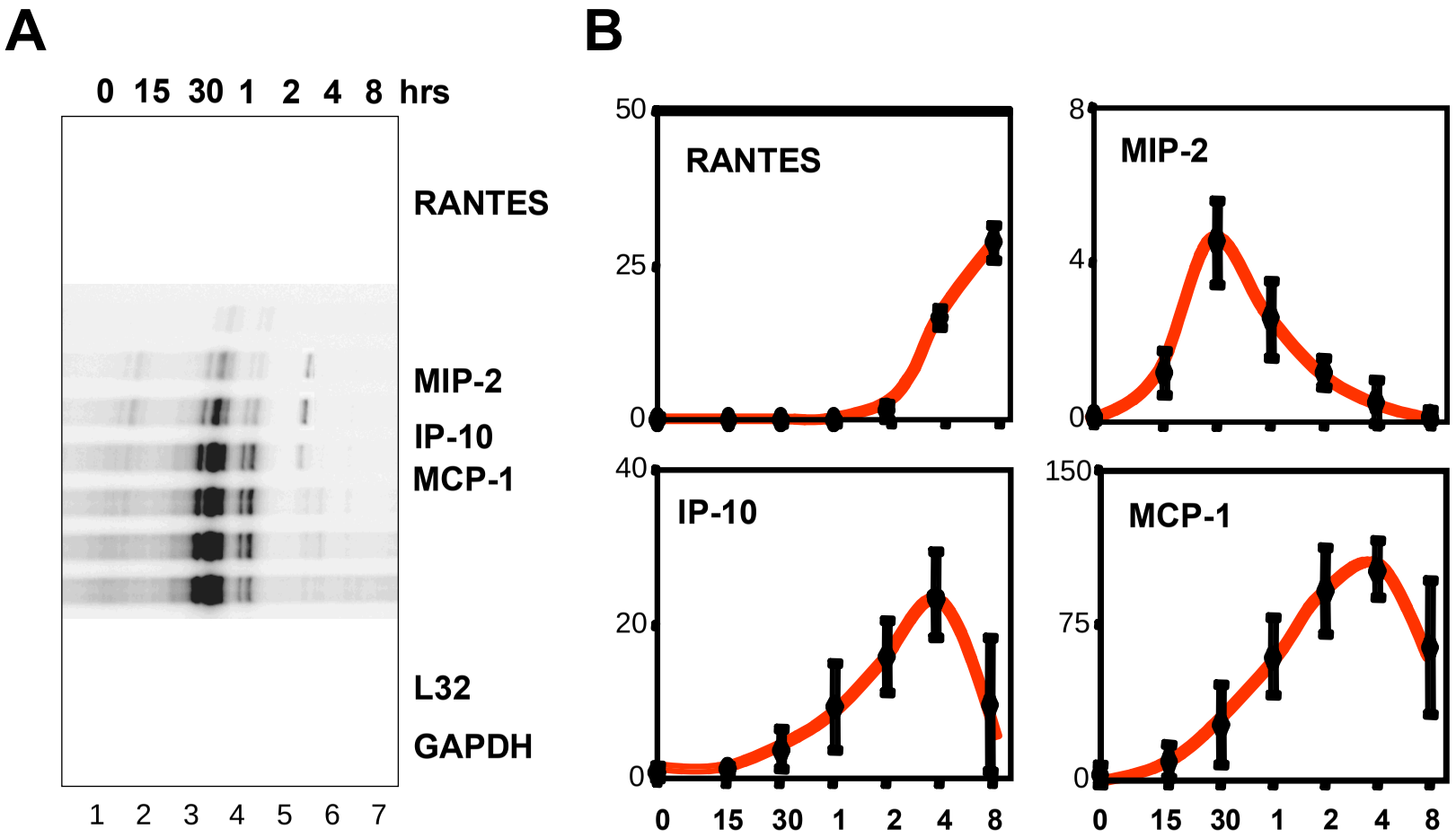
NF-κB-dependent genes (column 1) are listed in order of increasingly restrictive RHD protein dimer requirement (column 2). Corresponding gene promoters contain conserved κB-sites of indicated sequences (column 3) that were show to be functional in previously published studies (column 4).

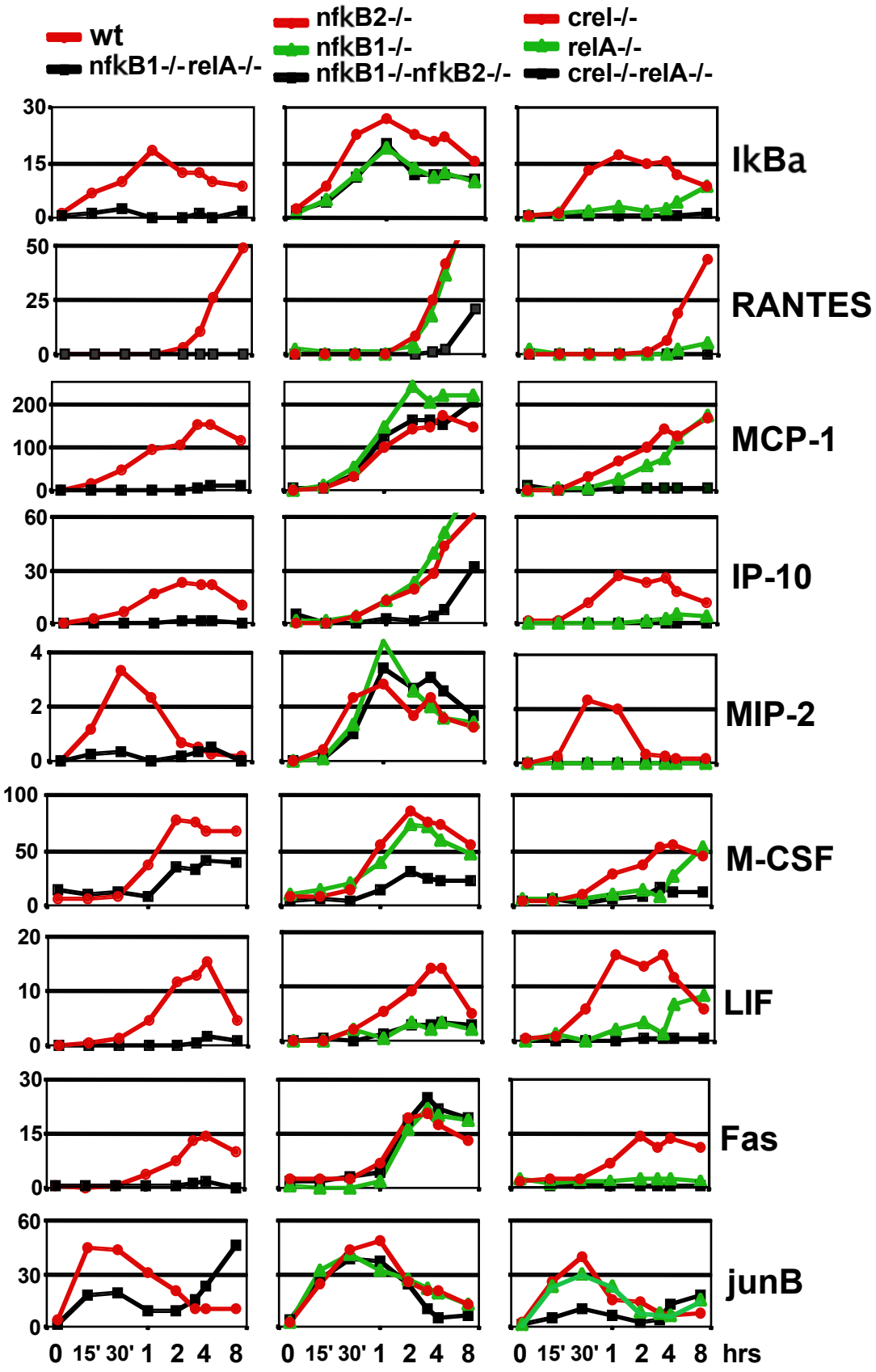

A
wt
nfkB1-/-
relA-/-
nfkB1-/-relA-/-
p50
p52
p65
cRel
RelB
actin
1 2 3 4
B
wt
nfkB1-/-
relA-/-
nfkB1-/-relA-/-
IkB-e
IkB-b
IkB-a
actin
1 2 3 4


wt
nfkB1-/-
relA-/-
nfkB1-/-relA-/-
A
C
E
G
0 10' 30' 1 90' 2 3 4 6 hrs
1 2 3 4 5 6 7 8 9
B
D
F
e
b
a


A
wt
comp.
antibody
–
wt
mt
p50
p52
p65
cRel
RelB
all
*
1 2 3 4 5 6 7 8 9
B
nfkB1-/-
comp.
antibody
–
wt
mt
p50
p52
p65
cRel
RelB
all
*
1 2 3 4 5 6 7 8 9
C
relA-/-
comp.
antibody
–
wt
mt
p50
p52
p65
cRel
RelB
all
*
*
1 2 3 4 5 6 7 8 9
D
nfkB1-/-relA-/-
comp.
antibody
–
wt
mt
p50
p52
p65
cRel
RelB
all
*
*
1 2 3 4 5 6 7 8 9


A
Igk-
H2-
IFNb-
G
20
15
10
5
0
0 2 4 8 12 hr
B
wt
nfkB1-/-
relA-/-
nfκB1-/-
relA-/-
0 .5 4 0 .5 4 0 .5 4 0 .5 4 hr
GGGACTTTCC
Igk
GGGATTCCCC
H2
GGGAAATTCC
IFNb
p50/p65
p65/p65
p52/p65
cRel/p50
p50/p50
–


A
0 15 30 1 2 4 8 hrs
RANTES
MIP-2
IP-10
MCP-1
L32
GAPDH
1 2 3 4 5 6 7
B
RANTES
MIP-2
IP-10
MCP-1
50
25
0
8
4
0
40
20
0
150
75
0
0 15 30 1 2 4 8


wt
nfkB1-/-relA-/-
nfkB2-/-
nfkB1-/-
nfkB1-/-nfkB2-/-
crel-/-
relA-/-
crel-/-relA-/-
IkBa
RANTES
MCP-1
IP-10
MIP-2
M-CSF
LIF
Fas
junB
0 15' 30' 1 2 4 8
0 15' 30' 1 2 4 8
0 15' 30' 1 2 4 8 hrs

MEFs genotype	κB-binding proteins				κB-binding complexes in EMSA		
	p50	p52	p65	cRel			
wt	+	–	+	–	p50/p50*	p50/p65	p65/p65*
nfκB1-/-	–	+	+	–		p52/p65	p65/p65*
nfκB1-/-nfκB2-/-	–	–	+	–			p65/p65
nfκB2-/-	+	–	+	–	p50/p50*	p50/p65	p65/p65*
relA-/-	+	–	–	+	p50/p50	p50/cRel	
relA-/-crel-/-	+	–	–	–	p50/p50		
crel-/-	+	–	+	–	p50/p50*	p50/p65	p65/p65*
nfκB1-/-relA-/-	–	–	–	–		–	

*detected with some but not all κB-site containing probes

	wt time	nfkB1-/- relA-/-	nfkB2-/-	nfkB1-/-	nfkB1-/- nfkB2-/-	relA-/- crel-/-	RHD protein required p50	p50 p52	p65	p65 cRel
IkBa	E	-	+	+	+	-	-	-	+	-
RANTES	L	-	+	+	-	-	-	+	+	-
MCP-1	M	-	+	+	+	-	-	-	-	+
IP-10	M	-	+	+	-	-	-	+	+	-
MIP-2	E	-	+	+	+	-	-	-	+	-
M-CSF	M	+/-	+	+	+/-	-	-	-	+	-
LIF	L	-	+	-	-	-	+	-	+	-
Fas	M	-	+	+	+	-	-	-	+	-
JunB	E	+/-	+	+	+	+/-	-	-	-	+
c-fos	E	+	+	+	+	+	-	-	-	-

gene	functioning RHD protein dimers				kB-site sequence		
MCP-1	p50/p65	p52/p65	p50/cRel	p65/p65	-2611 GGGAATTTCC	-2640 GGGAACTTCC	Ping *et al.* 1999
JunB	p50/p65	p52/p65	p50/cRel	p65/p65	+2057 GGGGCTTTCC		Phinney *et al.* 1995
MIP-2	p50/p65	p52/p65		p65/p65	-66 GGGAATTTCC		Widmer *et al.* 1993
IkBa	p50/p65	p52/p65		p65/p65	-260 GGGGAAGTCC	-70 GGAAATTCCC	Chiao *et al.* 1994
Fas	p50/p65	p52/p65		p65/p65	-59 GGAATGCCCA		Zheng *et al.* 2001
RANTES	p50/p65	p52/p65			-87 GGGAGTTTCC		Lee *et al.* 2000
IP-10	p50/p65	p52/p65			-169 GGGAAATTCC	-113 GGGACTTCCC	Ohmori *et al.* 1992
M-CSF	p50/p65	p52/p65			-378 GGAAAGTCCC		Harrington *et al.* 1991
LIF	p50/p65				-77 GGGGATCCCG		Willson *et al.* 1992

CHAPTER 4:

One Nucleotide in a κB Site Can Determine Cofactor Specificity for NF-κB Dimers

Thomas H. Leung, Alexander Hoffmann, and David Baltimore

Published in 2004 Cell 118: 453-464.

ABSTRACT

The transcription factor NF-κB regulates a wide variety of genes involved in multiple processes. Although the apparent consensus sequence of DNA-binding sites for NF-κB (κB sites) is very broad, the sites active in any one gene show remarkable evolutionary stability. Using a lentivirus based methodology for implantation of gene regulatory sequences we show that for genes with two κB sites, both are required for activity. Swapping sites between κB -dependent genes altered NF-κB dimer specificity of the promoters and revealed that two κB sites can function together as a module to regulate gene activation. Further, although the sequence of the κB site is important for determining κB family member specificity, rather than determining the ability of a particular dimer to bind effectively, the sequence affects which co-activators will form productive interactions with the bound κB dimer. This suggests that binding sites may impart a specific configuration to bound transcription factors.

INTRODUCTION

Study of mammalian gene transcription is often confounded by the recognition of a single regulatory sequence by multiple members of a transcription factor family. Expansion of the number of members in a particular family may occur over evolutionary time suggesting that particular members evolve to serve restricted functions in a complex organism. However, outside of the nuclear hormone receptor superfamily (Rastinejad, 2001), little is known about how individual family members provide specificity.

NF-κB is a dimeric protein that mediates cellular responses to a wide variety of stimuli including TNFα, LPS, IL-1, and ultraviolet light (Ghosh et al., 1998; Karin and Ben-Neriah, 2000; Karin et al., 2002). NF-κB plays an integral role in many important and diverse processes, including inflammatory responses, immune system development, apoptosis, learning in the brain, and bone development. In resting cells, NF-κB is held inactive via associations with inhibitory proteins called inhibitors of NF-κB (IκB). When the cell is stimulated, IκB proteins are phosphorylated, ubiquitinated, and degraded, allowing NF-κB to bind DNA and activate appropriate target genes.

In response to inflammatory stimuli, four members of the NF-κB family are involved in gene activation: p50, p52, p65, and cRel. They are the mature gene products from four genes: nfκB1, nfκB2, relA, and cRel respectively, and they homo- or heterodimerize with one another to bind DNA. It is noteworthy that mice deficient in a single NF-κB/Rel family member exhibit unique, largely non-overlapping phenotypes. For instance, relA$^{-/-}$ animals are embryonic lethal and develop massive liver apoptosis (Beg et al., 1995), while nfκB1$^{-/-}$ mice develop normally but have deficits in immune function (Sha et al., 1995).

Past studies predominantly used transient-transfection reporter systems with over-expressed NF-κB proteins to conclude that certain κB complexes bind particular NF-κB binding sites (called κB sites) preferentially (Fujita et al., 1992; Kunsch et al., 1992). When similar transient-transfection reporter assays were repeated in NF-κB knockout cells, NF-κB family members exhibited no preference for particular κB sequences (Hoffmann et al., 2003). With increased knowledge about transcriptional regulation, questions about the recruitment of specific NF-κB family members to a promoter should be addressed within a physiological context.

Our previous study utilized a genetic approach to study NF-κB-dependent gene activation by creating a panel of single and double NF-κB knockout cell lines (Hoffmann et al., 2003). We showed there that TNFα-stimulated, κB-dependent genes may require specific κB family members for activation. Some genes had stricter requirements than others. For example, any p65 or cRel-containing hetero- or homodimer could activate the MCP-1 gene, while the LIF gene specifically required a p50:p65 heterodimer for activation. We were unable to document a strict correlation between κB family member requirements and their κB site sequences.

Structural studies have also not revealed evidence for highly specific interactions between NF-κB and its cognate DNA-binding sites. κB sites display a remarkably loose consensus sequence, often cited as $G_{-5}G_{-4}G_{-3}R_{-2}N_{-1}N_0Y_{+1}Y_{+1}C_{+2}C_{+3}$ (Ghosh et al., 1998). Crystal structures have shown that p50 homodimers often interact specifically with 5'-$G_{-5}G_{-4}G_{-3}$, while p65 homodimers interact specifically with 5'-$G_{-5}G_{-4}$. X-ray structures of p50:p65 heterodimers demonstrate that they can bind to a variety of κB site sequences (Berkowitz et al., 2002; Chen and Ghosh, 1999; Chen-Park

et al., 2002; Escalante et al., 2002). Taken together, structural studies have not identified enough dimer-specific DNA-binding contacts to rationalize specific NF-κB binding sites and underscore the remarkable permissiveness in NF-κB-DNA interactions.

In this study, we demonstrate that the sequence of the κB site does play an important role in determining κB family member specificity. But rather than the site determining the ability of a particular dimer to bind effectively, we find that the sequence of the κB site affects which co-activators will form productive interactions with the bound κB dimer, suggesting that binding sites may affect the configuration of the bound dimer.

RESULTS

κB site sequences are strictly conserved between mouse and human

To understand whether the sequences of κB sites play a significant role in determining functional specificity, validated mouse and human κB site sequences for eleven κB-dependent genes were compared using the Celera and Ensembl databases (Hoffmann et al., 2003; Libermann and Baltimore, 1990; Thanos and Maniatis, 1995b). The κB site sequences for all eleven genes studied were 100% conserved between mouse and human (Figure 1A). In contrast, a pairwise sequence comparison program (Family Relations) revealed that the regions surrounding the κB sites often displayed <85% conservation (data not shown) (Brown et al., 2002). If the sequence of the κB site did not play a significant role in determining functional specificity, we would have expected the sequence of some of the κB sites to mutate over time. We therefore decided to test whether the sequence of the κB site can alter the functional properties of bound dimers.

To address the role of κB site sequence, we wanted to swap the sequences between two κB-dependent genes with different properties. The IP-10 and MCP-1 genes were chosen for investigation. When IP-10 and MCP-1 respond to TNF□, they have different κB family member requirements. MCP-1 is induced in both wild type (WT) and p50/p52 deficient cells, while IP-10 is induced in WT cells but not p50/p52 deficient cells (Hoffmann et al., 2003). p65 homodimers are the only detectable dimer in p50/p52 deficient 3T3 cells. Therefore, both hetero- and homodimers of NF-κB are able to function on the MCP-1 promoter while only heterodimers of NF-κB appear able to activate IP-10. The regulatory sequences of the genes have certain similarities but significant differences. Based on previous work as well as mouse/human sequence comparisons, MCP-1 and IP-10 each appear to have two functioning κB sites that are both required for gene activation (Ohmori and Hamilton, 1993; Ohmori and Hamilton, 1995; Ping et al., 1999). Both genes are highly expressed in fibroblasts and are responsive to multiple stimuli. However, the promoter architecture for the two genes is strikingly different. IP-10's κB sites are less than 200bp from the transcription start site while MCP-1's functional κB sites are more than 2.3kb away from the transcription start site. (MCP-1 has another apparent site near the promoter, which was shown to be non-functional and not conserved between mouse and human.) IRF-3 has been shown to be important in IP-10 gene regulation but not for MCP-1 (Sakaguchi et al., 2003). The κB site sequences are not dramatically different between MCP-1 and IP-10. All sites are 10bp long. The distal sites differ in one base pair, and the proximal sites differ in two base pairs (Figure1A).

Lentiviral-based reporter system recapitulates endogenous gene regulation

To effect the sequence interchanges, we needed a system that would allow us easy access for nucleotide alteration but would retain the regulatory properties of the endogenous gene promoters.

The use of a retrovirus vector allows insertion of up to 7kb of regulatory information and permits sequence alteration at will. We chose a self-inactivating lentiviral vector system that infects both dividing and non-dividing cells (Lois et al., 2002). The self-inactivating characteristic involves deletion of the virus's own promoter, ensuring that the inserted regulatory sequence is the only such information in the vector. We used a luciferase reporter gene to monitor transcriptional activity.

To test the reporter system, we cloned 5kb of the mouse MCP-1 gene promoter into the lentiviral construct. (We designate this ML M1M2, denoting its origin from the MCP-1 gene, its use of a luciferase reporter gene, and its two κB sites containing MCP-1-derived sequence). Lentiviral stocks were prepared, concentrated by ultracentrifugation, and applied to either wild type (WT) mouse 3T3 cells or 3T3 cells lacking both p50 and p65 subunits (effectively NF-κB-null cells (Hoffmann et al., 2003)). Routinely, multiple cell populations were created by infecting cells with serial dilutions of virus. Populations that expressed low basal luciferase activity (roughly 5-fold above background) were used for analysis to ensure a low number of integrated proviruses per cell.

WT and p50/p65 deficient cells with the integrated ML transgene were stimulated with TNFα (Figure 1B) over a 4 hour time course. WT cells responded robustly, while p50/p65 deficient cells had minimal inducible luciferase activity, showing that the system recapitulates the NF-κB dependence shown previously for the *in situ* gene. Next, the κB sites within the MCP-1 promoter were mutated to null κB site sequences (ML N1N2). WT cells carrying the integrated ML N1N2 transgene showed no inducible luciferase activity following TNFα treatment (Figure 1C). Finally, two different constructs were created with one κB site left intact and the other κB site mutated to a

null sequence (ML M1N2 and ML N1M2) (Figure 1D). As previously shown (Ping et al., 1999), the mutation of either κB site to a null sequence abrogated inducible luciferase activity in WT cells. These transgenic cells lines were also stimulated with LPS, and the results were similar to those with TNFα (data not shown). Thus the ML M1M2 transgene is inducible by multiple stimuli in a κB-dependent fashion, and both κB sites must be intact for the promoter to be functional.

Swapping MCP-1 κB site sequences for IP-10 κB site sequences imposes IP-10 κB family member requirements on the MCP-1 promoter

The κB site sequences from the IP-10 promoter were swapped into the MCP-1 transgene by PCR to create ML I1I2. WT cells infected with ML I1I2 were responsive to TNFα stimulation, but TNFα̃stimulated, ML I1I2-infected, p50/p52 deficient cells displayed no inducible luciferase activity (Figure 2A, right panel). In contrast, p50/p52 deficient cells infected with ML M1M2 displayed inducible luciferase activity comparable to that of WT cells (Figure 2A, left panel). We conclude that the ML I1I2 transgene requires NF-κB heterodimers for activation.

To test whether the altered κB family member requirements in ML I1I2 were dependent on one or both of the IP-10 κB site sequences, two MCP-1 promoter constructs were created carrying one IP-10 κB site and one MCP-1 κB site (ML M1I2, ML I1M2, Figure 2B). Both WT and p50/p52 deficient cells carrying these constructs remained responsive to TNFα stimulation. Therefore, either MCP-1 sequence alone is sufficient to provide a response to the p65 homodimer.

To determine if one or both of the IP-10 κB site sequences were unresponsive to the homodimer, two different MCP-1 promoter constructs containing duplicated IP-10 κB site sequences were

made (ML I1I1, ML I2I2, Figure 2C). WT cells infected with either ML transgene remained responsive to TNFα stimulation. However, p50/p52 deficient cells carrying the ML I1I1 transgene were not responsive to TNFα stimulation, while those carrying the ML I2I2 transgene displayed inducible luciferase activity. Thus, the I1 or distal κB site of the IP-10 promoter determines the heterodimer requirement of the IP-10 promoter. The κB site sequences in the ML I1I1 construct differ from those in the ML M1M2 construct only at the sixth nucleotide for each site. Therefore, we can identify this single nucleotide as responsible for the difference in κB family member requirements.

To exclude the possibility that nearby factors bound to the MCP-1 promoter may affect κB-dimer specificity, lentiviral constructs containing only the c-fos minimum promoter and the κB sites from MCP-1 or IP-10 were tested. Upon TNFα treatment, the construct containing the MCP-1 κB sites was responsive to hetero- and homodimers of NF-κB, while the construct containing IP-10 κB sites only responded to heterodimers of NF-κB (data not shown).

p65 and CBP are bound to the inactive ML I1I2 transgene in TNFα̃stimulated p50/p52 deficient cells

One possible explanation of the inability of the IP-10 distal κB site to respond to a homodimer of p65 would be an inability of the site to bind the homodimers. To address this possibility, we performed chromatin immunoprecipitation (ChIP) assays with an anti-p65 antibody in TNFα-stimulated WT and p50/p52 deficient cells containing the ML I1I2 transgene. WT cells with the ML I1I2 transgene showed robust p65 binding after TNFα stimulation (Figure 3A). Surprisingly, p50/p52 deficient cells with the ML I1I2 transgene also demonstrated p65 binding after TNFα

stimulation. Therefore, even though the ML I1I2 transgene shows minimal activity in TNFα stimulated p50/p52 deficient cells, p65 homodimers are still bound to the promoter. Examination of the endogenous IP-10 promoter gave a congruent result. In TNFα-stimulated p50/p52 deficient cells, p65 was still bound to an inactive IP-10 promoter (Figure 3B).

If p65 homodimers are bound to an inactive ML I1I2 transgene, the κB site sequence might not support transcription because the conformation of the p65 homodimers might not recruit the appropriate co-activators to the promoter. One type of co-activator known to interact with the transactivation domain of p65 is CBP/p300 (Perkins et al., 1997; Sheppard et al., 1999). We performed ChIP assays against p300 using TNFα-stimulated, ML M1M2-infected, p50/p52 deficient cells and discovered that p300 is recruited to the MCP-1 and IP-10 promoters in a stimulus-dependent manner (Figure 3C). Furthermore, we found that p300 was bound to the ML I1I2 transgene in TNFα stimulated p50/p52 deficient cells (Figure 3D). Thus, this co-factor is recruited even though the promoter is not firing.

IP-10 displays stimulus-specific requirements for NF-κB family members

To test whether the IP-10 gene shows the same heterodimer requirement to an inducer other than TNFα, WT and p50/p52 deficient cells were stimulated with TNFα or LPS over a 6 hour time course, and IP-10 induction was assayed by quantitative PCR (Q-PCR). The induction of IP-10 in TNFα̃stimulated WT and p50/p52 deficient cells recapitulated our previously published ribonuclease protection assays showing the heterodimer requirement for activation (Figure 4A, left panel). In contrast, LPS-stimulated WT and p50/p52 deficient cells induced IP-10 to similar levels (Figure 4A, right panel). Thus the heterodimer requirement is an inducer-specific phenomenon.

To determine if the lentiviral-based reporter system could recapitulate IP-10's stimulus-specificity, 1kb of the IP-10 promoter was cloned into our retrovirus vector (IL). Consistent with the whole cell data, TNFα-stimulated WT cells with the integrated IL transgene displayed induced luciferase activity, while p50/p52 deficient cells with the IL transgene showed significantly lower activity (Figure 4B, left panel). In contrast, LPS-stimulated WT and p50/p52 deficient cells with the IL transgene demonstrated similar luciferase activity (Figure 4B, right panel). Thus, the lentiviral system recapitulated the stimulus-specific behavior of the endogenous IP-10 gene.

To be certain that the stimulus-specificity was determined by the κB site, we went back to the chimeric transgenes. In fact, as opposed to the data with TNFα-stimulated cells, LPS-treated WT and p50/p52 deficient cells containing the ML I1I2 transgene displayed similar levels of luciferase induction (Figure 4C, right panel). Studies with the ML I1I1 transgene demonstrated similar results (data not shown) and supported the conclusion that p65 homodimers are bound to the I1 site. Taken together, these results confirmed that the stimulus-specificity was a consequence of the κB site sequence and was consistent with the observation of p65 homodimers on the ML I1I2 transgene. It strongly suggested that the unresponsiveness of the IP-10 gene to homodimers in the TNFα̃treated cells is a consequence of the lack of a co-factor that LPS can induce.

Overexpression of IRF-3 rescues induction of ML I1I2 transgene by TNFα

To identify the LPS-specific factor that allows κB homodimers to activate the ML I1I2 transgene, we noted that LPS signals through the TLR-4 receptor (Akira, 2003; Takeda et al., 2003). TLR-4 downstream signaling is divided into two distinct pathways, one mediated by Myd88 and the other

by Trif. Since IP-10 is induced by LPS in Myd88 knockout fibroblasts (Yamamoto et al., 2002), we focused on signaling molecules in the Trif-dependent pathway. IRF-3 is a transcription factor induced in the Trif-dependent pathway and has been shown to play a role in IP-10 induction by LPS (Sakaguchi et al., 2003; Yamamoto et al., 2003). Furthermore, an interaction between IRF-3 and p65 has been demonstrated (Wietek et al., 2003).

A retrovirus expressing the IRF-3 cDNA was created and used to infect WT and p50/p52 deficient cells. Overexpression of IRF-3 protein was confirmed by Western analysis (Figure 5A, left panel). IRF-3 overexpressing cells were then infected with the ML I1I2 transgene and stimulated with TNFα over a 6 hour time course. Overexpression of IRF-3 rescued the induction of the ML I1I2 transgene in p50/p52 deficient cells (Figure 5A, right panel). This result also supports the interpretation of the ChIP data that p65 was bound to the identified functional κB sites and not cryptic elements.

Since the MCP-1 promoter does not contain an IRF-3 binding site, it seemed likely that IRF-3 is serving as a co-activator in this situation. Previous work has shown that the c-myc transcription factor can bind and activate a number of promoters, which contain no apparent c-myc binding sites (Li et al., 2003). To test this hypothesis, an expression vector containing IRF-3 with its DNA-binding domain deleted was created (IRF-3ΔDBD). Expression of IRF-3 ΔDBD in WT and p50/p52 deficient cells was confirmed by Western analysis (Figure 5B, left panel). The IRF-3 ΔDBD cells were then infected with the ML I1I2 transgene and stimulated with TNFα over a 6 hour time course. Expression of IRF-3 ΔDBD rescued the induction of the ML I1I2 transgene in p50/p52 deficient cells (Figure 5B, right panel).

To confirm that IRF-3 is required for LPS stimulation of the ML I1I2 transgene in p50/p52 deficient cells, we created a retrovirus expressing an RNAi cassette against IRF-3 (Qin et al., 2003). IRF-3 protein levels were confirmed by Western analysis and were knocked down approximately 3-4-fold (Figure 5C, left panel). p50/p52 deficient cells with IRF-3 knocked-down were infected with the ML I1I2 transgene, stimulated with LPS, and assayed for luciferase induction. The absence of IRF-3 in LPS-stimulated p50/p52 deficient cells abrogated the induction of the ML I1I2 transgene (Figure 5C, right panel).

As mentioned earlier, endogenous IP-10 requires IRF-3 for gene activation. We have demonstrated that the κB site sequence can alter the configuration of bound p65:p65 homodimers to implement an IRF-3 requirement for transgene activation. Does the change in κB site sequence alter the configuration of all κB dimers such that the ML I1I2 transgene requires IRF-3 for gene activation in WT cells, where the major species of NF-κB is p65:p50 heterodimers? Using the same retrovirus, we knocked down IRF-3 in WT cells (Figure 5D, left panel). The absence of IRF-3 in LPS-stimulated WT cells also abrogated the induction of the ML I1I2 transgene (Figure 5D, right panel). To further confirm these findings, the absence of IRF-3 in TNFα-stimulated WT cells did not affect the induction of IP-10 or the ML I1I2 transgene (data not shown).

IRF-3 is recruited to I1I2 κB-binding sites

To test whether IRF-3 is directly recruited to the ML I1I2 transgene, ChIP assays were performed with an anti-IRF-3 antibody in LPS-stimulated, IRF-3 overexpressing, WT cells. We found that IRF-3 was bound to the IP-10 promoter but not the MCP-1 promoter in a LPS-dependent manner

(Figure 5E, left panel). Furthermore, IRF-3 was also bound to the ML I1I2 transgene (Figure 5E, right panel). Therefore, IP-10's κB site sequences are able to recruit IRF-3 to a promoter.

IP-10's requirement for IRF-3 in p50:p52 deficient cells can explain the stimulus-specific requirements for κB family members. Under LPS stimulation, p65 homodimers appear to interact with IRF-3 to activate the ML I1I2 transgene. Under TNFα stimulation, IRF-3 is not activated and the p65 homodimers, although bound, are unable to function and the ML I1I2 transgene is not activated. For LPS stimulated WT cells, we demonstrated that p65:p50 heterodimers appear to interact with IRF-3 to activate the ML I1I2 transgene. This result implies that the κB site sequence can change the configuration of heterodimers, not just p65 homodimers. Finally, we show that IRF-3 is directly recruited to the ML I1I2 promoter in a stimulus-dependent manner.

IP-10 requires Bcl-3 for Induction by TNFα

The need for IRF-3 provides an explanation for the stimulus specificity in p50/p52 deficient cells. But we see no such specificity in WT cells. Why not? Might there be another co-activator that suffices in WT cells where the p65:p50 heterodimer predominates? We were drawn to the observation that Bcl-3, an IκB-like protein that interacts preferably with p50 or p52, can serve as a co-activator (Franzoso et al., 1992; Fujita et al., 1993).

To examine the possibility that Bcl-3 is critical to IP-10 induction in WT cells, Bcl-3-deficient cells were stimulated with TNFα over a 4 hour time course. RNA transcript levels of endogenous IP-10 and MCP-1 were measured by Q-PCR. IP-10 induction was absent in Bcl-3 deficient cells, while MCP-1 induction recapitulated WT levels (Figure 6A). Thus, in TNFα-stimulated cells, IP-

10 is a Bcl-3 dependent gene. As expected, when Bcl-3 deficient cells were infected with the ML I1I2 transgene and stimulated with TNFα over a 4 hour time course, minimal luciferase induction was observed, showing that the Bcl-3 dependence is a consequence of the κB site sequence (Figure 6B, left panel). In contrast, the ML M1M2 transgene was induced to similar levels in WT and Bcl-3 deficient cells (Figure 6B, right panel). Finally, induction of IP-10 expression in LPS-stimulated Bcl-3 deficient cells remained intact (data not shown).

DISCUSSION

In our previous study, no direct correlation between the κB site sequence and κB family member requirements for gene activation could be found (Hoffmann et al., 2003). There we compared one gene to another but when, in this study, we made interspecies comparisons of the same gene, we found a remarkable constancy of sequence, implying that the individual sequences have important characteristics. This led us to examine the role of the particular sequences found associated with particular genes. To do this we developed a lentiviral system for incorporating regulatory sequences into cellular DNA. Then, by swapping the κB site sequences within the MCP-1 promoter to the κB site sequences for the IP-10 gene, we found that we could impose IP-10's κB family member requirements onto the MCP-1 promoter. Both IP-10 κB site sequences had to be transferred to change κB family member requirements and revealed that two κB sites can function together as a module to regulate gene activation. This suggested that either MCP-1 site was dominant over the two IP-10 sites. By doing chromatin immunoprecipitation experiments, we found that even though the I1 site would not work with the p65 homodimer, the IP-10 κB sites did bind the homodimer which, in turn, even bound the co-activator CBP/p300. We then found that the IP-10 requirement for a κB heterodimer for activation by TNFα is not evident after LPS

stimulation. This suggested that the κB site specificity operated by imposing on the DNA-bound NF-κB a co-factor requirement for activation. In fact, we showed that the ML I1I2 transgene requires Bcl-3 when stimulated by TNFα and that IRF-3 can play this role in p50/p52 deficient cells where Bcl-3 is not able to function. Because IRF-3 is induced by LPS but not TNFα, the role of this κB site specificity is explained. We extended the analysis to show that the IRF-3 requirement for LPS-stimulated ML I1I2 transgene activity applies to p50:p65 heterodimers as well. Therefore, the κB site sequence affects the configuration of both heterodimers and p65 homodimers. Finally, we showed that IRF-3 is recruited to the ML I1I2 promoter. The sequence alteration that imposes the co-activator requirement is a single nucleotide in the sixth position of the κB site—it is quite remarkable that such a change can impose on a gene a new co-factor requirement that is fulfilled only under particular circumstances.

How is specificity imposed?

There are three models for how the single nucleotide difference in the κB site can impose such specificity. One model suggests that there is another protein bound to the DNA site that requires this particular nucleotide. It would most likely bind in conjunction with NF-κB, just as HMG I(Y) has been shown to bind to certain κB sites along with NF-κB (Falvo et al., 1995). The factor would bind the distal IP-10 site because that site dominantly imposes the heterodimer restriction. It seems a bit far-fetched but not impossible that a protein could bind to the κB site and then dominantly impose a restriction on the functioning of a dimer that would be overcome by a co-factor. IRF-3 and Bcl-3 would be the responsible co-factors in LPS- and TNFα-stimulated cells respectively.

Another possibility is that κB dimers may oligomerize when bound to their respective sites. The sequence of the κB site would determine oligomerization efficiency and precise conformation of the overall structure. The κB tetramer would determine which co-factors would be needed for gene activation.

A likely possibility is that the particular distal IP-10 κB site imposes a configuration on the bound heterodimers that establishes a requirement for Bcl-3. In the p50/p52 deficient cells, where Bcl-3 cannot bind the p65 homodimers, the requirement for a co-activator can be supplied by IRF-3, which can bind to p65 (Wietek et al., 2003).

Potential model

To fully explain our data, we return to the observation that two κB sites are needed for the MCP-1 promoter to function. The stimulus and subunit specificities have been explained by the role of one κB site, but why then do we need two? We suspect that the two sites serve different and non-redundant roles. This postulate leads us to a model that the two κB sites serve different and jointly obligate functions.

In LPS-stimulated p50/p52 deficient cells infected with the ML M1M2 transgene, we know that p65 homodimers are bound to the κB sites along with p300/CBP (data not shown), and drive luciferase expression. Our postulate that the two sites serve different functions leads us to suppose that only one κB homodimer binds p300 and that the other binds an unknown factor X but other explanations are conceivable (Figure 7A). In LPS-stimulated p50/p52 deficient cells infected with the ML I1I2 transgene, we postulate that the I2 site continues to bind p300/CBP but the I1 κB site

sequence alters the p65 homodimer conformation such that another co-factor must take the place of X. IRF-3 can serve this function and in LPS-induced cells, it binds and cooperates to drive luciferase expression (Figure 7B). In TNFα-stimulated p50/p52 deficient cells with the ML M1M2 transgene, the situation is similar to that in LPS-simulated cells (Figure 7C). However, in the TNFα-stimulated p50/p52 deficient cells with the ML I1I2 transgene, the change in the I1 κB site sequence alters the conformation for p65 homodimers such that they require a co-factor of a type that is simply not present in the cells. Instead, only κB heterodimers are able to drive transcription because of the requirement for the second activity is supplied by Bcl-3 binding to the p50 subunit (Figure 7D).

Advantages of a retrovirus-based reporter system

Our retrovirus-based lentiviral reporter system is a novel approach to study gene transcription. This system offers many advantages over current transcription reporter systems. 1. The promoter of the gene of interest is integrated into the cell's DNA and should therefore be subject to chromatin regulation. 2. Tandem copies of integration are avoided to allow for faithful promoter regulation. 3. Copy number of the transgene can be controlled. 4. This system can be applied to non-dividing and untransfectable cells. 5. Most importantly, the system allows flexibility to conveniently modify promoter sequence by PCR. A limitation of this system is the amount of DNA sequence one can place into the vector (roughly 7kB). In truth, it is not possible to know for sure how much regulatory DNA a particular gene requires so the optimum method for studying promoters would be to make nucleotide changes *in situ* in the promoter regulatory elements using "knock-in" technology (Xu et al., 1996). Unfortunately, "knock-in" or even BAC transgene experiments (Lee et al., 2003) are difficult, time-consuming and prohibitively expensive as a method to do a

systematic study, such as we have presented here. Our approach represents an effective compromise between the ideal and need for extensive experimental manipulation.

κB site sequence determines NF-κB family member requirements

In our previous study, we documented that κB-dependent genes require specific κB family members for functional activation (Hoffmann et al., 2003). Here, we demonstrate that the sequence of the κB site is responsible for determining NF-κB dimer specificity, and two κB binding sites can function together to regulate gene activation. This regulation does not occur at the level of κB dimer exclusion. In fact, we found inducible p65 recruitment to a functionally inactive promoter. Instead, we believe that the sequence of the κB site contains information that is interpreted by the bound κB dimer, changes the κB dimer configuration, and determines which co-activators will form functional interactions with the κB dimer.

How many functionally different κB sites might there be? We can presently identify three different types. M1 or M2 are sites that, as far as we can determine are identical and which we postulate bind factor "X." A key property of this class of κB site is that it is dominant over other sites. The second class is I1, a site that requires IRF-3 or Bcl-3 co-activators for activation when it is coupled to I2 or is used in tandem. I1 differs from M1 and M2 at only one position, nucleotide 6. Finally, the I2 site is in a category by itself. It must require a co-activator different from M1 or M2 that is not dominant over I1. It is notable that I2 has two nucleotides, at positions 5 and 8, different from both M2 and I1. There is every reason to expect that there are other κB sites with specificities yet to be identified. From our limited evolutionary survey in Figure 1A, we see 12 different sites all of which are conserved between mouse and human. It could be that these 12 and others are all serving

specific functions, going a long way to providing the breadth of diversity needed to understand how κB sites can be found in so many different genes which are regulated in specific ways (Hoffmann et al., 2003).

κB site sequence determines co-activator requirements

It is impressive that changing one base pair would have such a dramatic effect on gene regulation. However, recent structural studies support this conclusion. Comparisons of p50:p65 dimers bound to different κB sites sequences have revealed significant differences in their conformation (Chen-Park et al., 2002). These differences in conformation have been correlated to their ability to transactivate transfected reporter plasmids. Therefore, our notion that it is co-activator binding which determines specificity is supported by prior structural studies showing alternate conformations induced allosterically by different DNA-binding site sequences.

In other families of transcription factors, there is also evidence that DNA can act as an allosteric regulator. It has been best-studied with the glucocorticoid receptor (GR) (Lefstin and Yamamoto, 1998). The DNA-binding domain of GR can bind its cognate site in a variety of conformations depending on the sequence of the site, serving to present different interfaces for the recruitment of specific co-factors and to help determine whether the bound factor will activate or repress transcription. Recent work has identified which GR surfaces are required in a gene-specific manner (Rogatsky et al., 2002; Rogatsky et al., 2003). Similar findings have also been reported for the POU family of proteins (Phillips and Luisi, 2000).

The prevailing model of gene transcription is that on a given promoter multiple proteins must interact, assemble, and form an enhanceosome to activate gene transcription (Thanos and Maniatis, 1995b). If one factor is absent, the gene is not activated. Numerous examples of NF-κB's role in this model exist, most notably on the HIV-LTR, IFNß, and IL-2 promoters (Perkins et al., 1994; Perkins et al., 1993; Rothenberg and Ward, 1996; Thanos and Maniatis, 1995b). Furthermore, chromatin has also been implicated as having a role in determining NF-κB's access to DNA-binding sites and may do so in a κB dimer-specific fashion (Saccani et al., 2001; Saccani et al., 2003). In the best-studied system, the IFNß gene, there is only one κB site and that is why it may be so complex. *In vitro* binding and transient transfection studies demonstrated that the κB site sequence is important, presumably to coordinate the binding of both NF-κB and HMG (I)Y (Thanos and Maniatis, 1995a). As we imply by the models in Figure 7, where there are two κB sites, the situation may be different because there can be two interacting complexes formed around two NF-κB dimers.

Bcl-3 requirement

TNFα activation of IP-10 activation requires Bcl-3, and our data implies that p50:p65 heterodimers interact with Bcl-3. Previous studies have shown that Bcl-3 preferentially interacts with p50 or p52 homodimers and not with p50:p65 heterodimers (Franzoso et al., 1992; Fujita et al., 1993). We can suggest three potential mechanisms to explain this apparent discrepancy. First, even though the major constituent of NF-κB in wild type fibroblasts is p50:p65 heterodimers, a significant amount of p50:p50 homodimers exist and could bind one of the two sites to interact with Bcl-3 to activate IP-10. Second, it may be possible that p50:p65 heterodimers can interact with Bcl-3 to activate IP-10 expression. Third, as mentioned earlier, p50:p65 heterodimers may oligomerize when bound to

their respective sites. The p50 subunits from each heterodimer could conceivably interact and recruit Bcl-3,

Biology

MCP-1 and IP-10 are differentially regulated, presumably because they serve different functions. MCP-1 plays a significant role in innate immunity by bringing macrophages to sites of inflammation. On the other hand, IP-10 is important for lymphocytic (adaptive) immunity and regulates T cell proliferation. MCP-1 is also activated by a more diverse set of TLR agonists than IP-10. For example, TLR-2 agonists stimulate MCP-1, but not IP-10 expression (Pierer et al., 2004; Re and Strominger, 2001). Recent studies reveal that MCP-1 and IP-10 are regulated differently during TLR-4 stimulation (Sakaguchi et al., 2003; Serbina et al., 2003; Yamamoto et al., 2003). IP-10 is activated through TLR-4 via a Trif dependent pathway. The Trif dependent pathway activates the interferon-response pathway and is responsible for a late NF-κB activation (Yamamoto et al., 2003). It has been shown that this pathway regulates a specific subset of TLR3/TLR4 dependent genes, and that the pathway is evolutionarily diverged from other members of the TLR family (Doyle et al., 2002). In contrast, MCP-1 is activated by both the MyD88-dependent and Trif-dependent pathways. The Myd88 pathway differs from the Trif pathway in two ways. First, it does not activate the interferon-response pathway. Second, it activates NF-κB much earlier than the Trif pathway.

These two pathways normally work together to ready an immune response to a bacterial pathogen. A bacterial pathogen would signal the TLR-4 receptor and activate both downstream pathways. The Myd88 pathway would activate NF-κB immediately to drive MCP-1 production and recruit

macrophages to the target site. Later, the Trif pathway would activate both the NF-κB and IRF-3 pathways to drive IP-10 production and regulate T cell proliferation. By simply changing the sequence of κB sites in the MCP-1 promoter, we converted MCP-1's regulation profile into IP-10's. We speculate that in a mouse this would severely disrupt the delicate balance between innate and adaptive immunity. Our results underscore the functional importance of the sequence of the κB site and confirm why the sequence of κB sites is strictly conserved over time. Not only does the κB site sequence determine κB dimer specificity, it also determines co-activator requirements.

MATERIALS AND METHODS

Cell culture and Time Courses

Immortalized cell lines were generated and maintained as previous described (Hoffmann et al., 2003). 100% confluent cells were treated for 48 hours with 0.5% serum containing medium and then stimulated with the appropriate concentration of TNFα or LPS over a period of 4-6 hours.

Plasmids

To generate ML M1M2, 5kB of the MCP-1 promoter was amplified by PCR from a BAC clone. The resulting fragment was cloned into FugW (Lois et al., 2002). To generate IL, 967bp of IP-10 promoter (-939 - +28) was amplified by PCR and cloned into FugW. Variations of ML plasmid were constructed by PCR mutagenesis. IRF-3 and IRF-3 ΔDBD expression plasmids were generated by amplifying IRF-3 (amino acids 1-420) and IRF-3 ΔDBD (133-420) by PCR and cloning into the pBABE-neo vector (Morgenstern and Land, 1990). All plasmids were verified by DNA sequencing and restriction digest analysis.

Real Time PCR

Total RNA was made from confluent and starved fibroblasts using TriReagent (Molecular Research Center, Inc.). cDNA was synthesized with Superscript II (Invitrogen) following manufacturer's guidelines and 100ng total RNA as a starting amount. Q-PCR was performed using Sybr Green PCR Master Mix (Applied Biosystems) and a 5600 Real Time PCR machine (Applied Biosystems). Samples were performed in triplicates, and GAPDH transcript levels were used to

normalize between samples. Every experiment was performed at least twice, many three-six times with high reproducibility. Primer sequences are available upon request.

Reporter Assay

Luciferase assay was performed as described elsewhere (Pomerantz and Baltimore, 1999). Samples were normalized for protein levels by Bradford assay. Every experiment was performed at least twice, most three-five times with high reproducibility.

Chromatin Immunoprecipitation

ChIP studies were based on a combination of published protocols (Boyd and Farnham, 1999; Boyd et al., 1998; Nissen and Yamamoto, 2000; Saccani et al., 2001). p65 and IRF-3 antibodies were purchased from Santa Cruz Biotechnologies (sc-109) and Zymed Laboratories respectively. p300 antibody was a gift from Kevin Gardner. Sequences of promoter-specific primers and our detailed protocol are available upon request.

Lentivirus

Lentivirus was produced as described elsewhere (Lois et al., 2002).

siRNA

FG12 RNAi vector and production of siRNA was previously described elsewhere (Qin et al., 2003). The IRF-3-siRNA contains the sense targeting sequence of gacgcacagatggctgact corresponding to the 390-407 nucleotide positions of the mouse IRF-3 coding sequence.

Acknowledgments

We thank: E. Schwarz, K. Gardner, G. Cheng, and S. Tronick for providing reagents, G. Natoli for technical assistance, and M. Simon for use of the Real Time PCR machine. M. Boldin, E. Brown, H. Guo, C. Lois, W. Lu, J. Pomerantz and X. Qin for helpful advice and guidance. THL is a student in the UCLA-Caltech Medical Scientist Training Program and supported by the ARCS foundation. This study was supported by NIH grants RO1-A1042549-061 and RO1-GM39458-20.

FIGURE LEGENDS

1. Functional κB site sequences are strictly conserved and activation of lentiviral MCP-1 trangene is responsive to stimuli and dependent on NF-κB

Validated mouse and human κB site sequences for eleven κB-dependent genes were compared using the Celera and Ensembl databases (A). The κB site sequences for all of the eleven genes studied were 100% conserved between mouse and human. Roughly 5kb of the MCP-1 gene promoter and a luciferase reporter gene (ML M1M2) was implanted into wild-type and p50/p65 knockout 3T3 cells by lentiviral infection. Cells were stimulated with TNFα over a four-hour time course. Four variations of the ML transgene are shown: a 5kB wild-type version "ML M1M2" (B), a version with both κB binding sites mutated to a null sequence "ML N1N2" (C), a version where the distal κB binding site is intact and the proximal site is mutated to a null sequence "ML M1N2"(D, left panel), and a version where the distal binding site is mutated to a null sequence and the proximal site is intact "ML N1M2" (D, right panel).

2. Swapping MCP-1 κB site sequences for IP-10 κB site sequences imposes IP-10 κB family member requirements on the MCP-1 promoter.

IP-10 κB site sequences (ML I1I2) were swapped for the MCP-1 κB site sequences in the ML transgene by PCR mutagenesis. Lentivirus was produced and applied to wild-type and p50/p52 knockout 3T3 cells. Panel A compares the activity of the integrated ML M1M2 transgene to the ML I1I2 transgene. Panel B shows the activity of the transgene when one MCP-1 κB binding site is swapped with its corresponding IP-10 κB binding site sequence (ML M1I2 or ML I1M2). Panel C compares the activity of the transgene when either one of the IP-10 κB sites are duplicated and

inserted into the transgene (ML I1I1 and ML I2I2). All cells were stimulated with TNFα over a 4 hour time course.

3. p65 homodimer and p300/CBP bind to ML I1I2.

Chromatin immunoprecipitation assays with antibodies against p65 (A, B) and p300 (C, D) demonstrate that p65 and p300 binds the MCP-1 promoter, the IP-10 promoter, and ML I1I2 transgene in a TNFα-dependent manner. The white line indicates where lanes irrelevant to the figure have been cropped (B).

4. IP-10 κB binding sites display stimulus-specific κB family member requirements.

Wild-type (WT) and p50/p52 knockout cells were stimulated with LPS. IP-10 RNA transcript was measured by quantitative PCR (A). 1kb of the IP-10 promoter was cloned into the lentiviral transfer vector (IL). WT and p50/p52 knockout cell lines with integrated IL were stimulated with TNFα or LPS (B). WT and p50/p52 knockout cells integrated with the ML I1I2 transgene were stimulated with TNFα or LPS (C).

5. IRF-3 is recruited to the ML I1I2 transgene and required for transgene activation

IRF-3 and IRF-3 ΔDBD were overexpressed in WT and p50/p52 deficient cells carrying the ML I1I2 transgene (Panels A, B respectively). IRF-3 overexpressed cells were stimulated with TNFα. IRF-3 was knocked-down by RNAi in p50/p52 deficient and WT cells carrying the ML I1I2 transgene (Panels C, D respectively). IRF-3 knocked-down cells were stimulated with LPS. Chromatin immunoprecipitation assays with antibodies against IRF-3 demonstrate that IRF-3 is recruited to the IP-10 promoter but not the MCP-1promoter in a LPS-dependent manner (E, left

panel). IRF-3 is recruited to the ML I1I2 transgene in a LPS-dependent manner (E, right panel). The white line indicates where lanes irrelevant to the figure have been cropped.

6. Induction of ML I1I2 by TNFα Requires Bcl-3

Bcl-3 KO and Wild-type (WT) cells are stimulated with TNFα or LPS over a 4 hour time course. Induction of endogenous IP-10 and MCP-1 RNA transcript levels are measured by quantitative PCR (A). Bcl-3 KO and WT cells carrying the ML I1I2 or ML M1M2 transgene are stimulated with TNFα over a 4 hour time course (B).

7. Two κB sites function together to determine NF-κB family member and co-activator specificity

A model depicting how two κB sites could coordinate as a single module to regulate gene activation. The model postulates that the two κB sites in a promoter/enhancer have separable functions that must be jointly provided from the module to support transcription. One is to bind p300/CBP, the other is to bind a second co-activator whose nature is dictated by the sequence of the κB site to which that NF-κB dimer is bound. One hypothetical co-activator is designated X, the others are known. The model is presented in the context of 4 conditions discussed in the text. In LPS-stimulated p50/p52 deficient cells, p65 homodimers bound to M1 and M2 recruit p300/CBP and X (A). In the same cells, the alteration to I1 does not allow X binding but does support IRF-3 binding (B). In TNFα-stimulated p50/p52 deficient cells infected with ML M1M2, the situation is similar to that in LPS-stimulated cells (C). However, in TNFα-stimulated p50/p52cells infected with ML I1I2, because IRF-3 is not activated, the p65 homodimers cannot bind a co-activator and gene activation does not occur (not shown). However, in wild-type cells, NF-κB heterodimers are

present and p50 is able to bind Bcl-3, providing the co-activator needed for the module to function (D).

References

Aaronson, S. A., and Todaro, G. J. (1968). Development of 3T3-like lines from Balb-c mouse embryo cultures: transformation susceptibility to SV40. *J. Cell. Physiol. 72*, 141-148.

Akira, S. (2003). Toll-like receptor signaling. *J. Biol. Chem. 278*, 38105-38108.

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

Berkowitz, B., Huang, D. B., Chen-Park, F. E., Sigler, P. B., and Ghosh, G. (2002). The x-ray crystal structure of the NF-kappa B p50.p65 heterodimer bound to the interferon beta -kappa B site. *J. Biol. Chem. 277*, 24694-24700.

Boyd, K. E., and Farnham, P. J. (1999). Coexamination of site-specific transcription factor binding and promoter activity in living cells. *Mol. Cell. Biol. 19*, 8393-8399.

Boyd, K. E., Wells, J., Gutman, J., Bartley, S. M., and Farnham, P. J. (1998). c-Myc target gene specificity is determined by a post-DNAbinding mechanism. *Proc. Natl. Acad. Sci. USA 95*, 13887-13892.

Bren, G. D., Solan, N. J., Miyoshi, H., Pennington, K. N., Pobst, L. J., and Paya, C. V. (2001). Transcription of the RelB gene is regulated by NF-kappaB. *Oncogene 20*, 7722-7733.

Brown, C. T., Rust, A. G., Clarke, P. J., Pan, Z., Schilstra, M. J., De Buysscher, T., Griffin, G., Wold, B. J., Cameron, R. A., Davidson, E. H., and Bolouri, H. (2002). New computational approaches for analysis of cis-regulatory networks. *Dev. Biol. 246*, 86-102.

Caamano, J. H., Rizzo, C. A., Durham, S. K., Barton, D. S., Raventos-Suarez, C., Snapper, C. M., and Bravo, R. (1998). Nuclear factor (NF)-kappa B2 (p100/p52) is required for normal splenic microarchitecture and B cell-mediated immune responses. *J. Exp. Med. 187*, 185-196.

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

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.

Chen-Park, F. E., Huang, D. B., Noro, B., Thanos, D., and Ghosh, G. (2002). The kappa B DNA sequence from the HIV long terminal repeat functions as an allosteric regulator of HIV transcription. *J. Biol. Chem. 277*, 24701-24708.

Cheng, J. D., Ryseck, R. P., Attar, R. M., Dambach, D., and Bravo, R. (1998). Functional redundancy of the nuclear factor kappa B inhibitors I kappa B alpha and I kappa B beta. *J. Exp. Med. 188*, 1055-1062.

Chinenov, Y., and Kerppola, T. K. (2001). Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. *Oncogene 20*, 2438-2452.

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

Escalante, C. R., Shen, L., Thanos, D., and Aggarwal, A. K. (2002). Structure of NF-kappaB p50/p65 heterodimer bound to the PRDII DNA element from the interferon-beta promoter. *Structure* (Camb) *10*, 383-391.

Falvo, J. V., Thanos, D., and Maniatis, T. (1995). Reversal of intrinsic DNA bends in the IFN beta gene enhancer by transcription factors and the architectural protein HMG I(Y). *Cell 83*, 1101-1111.

Falvo, J. V., Uglialoro, A. M., Brinkman, B. M., Merika, M., Parekh, B. S., Tsai, E. Y., King, H. C., Morielli, A. D., Peralta, E. G., Maniatis, T., et al. (2000). Stimulus-specific assembly of enhancer complexes on the tumor necrosis factor alpha gene promoter. *Mol. Cell. Biol. 20*, 2239-2247.

Franzoso, G., Bours, V., Park, S., Tomita-Yamaguchi, M., Kelly, K., and Siebenlist, U. (1992). The candidate oncoprotein Bcl-3 is an antagonist of p50/NF-kappa B-mediated inhibition. *Nature 359*, 339-342.

Fujita, T., Nolan, G. P., Ghosh, S., and Baltimore, D. (1992). Independent modes of transcriptional activation by the p50 and p65 subunits of NF-kappa B. *Genes Dev. 6*, 775-787.

Fujita, T., Nolan, G. P., Liou, H. C., Scott, M. L., and Baltimore, D. (1993). The candidate proto-oncogene bcl-3 encodes a transcriptional coactivator that activates through NF-kappa B p50 homodimers. *Genes Dev. 7*, 1354-1363.

Gerondakis, S., Grossmann, M., Nakamura, Y., Pohl, T., and Grumont, R. (1999). Genetic approaches in mice to understand Rel/NF-kB and IkB function: trangenics and knockouts. *Oncogene 18*, 6888-6895.

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.

Ghosh, S., May, M. J., and Kopp, E. B. (1998). NF-kB and Rel proteins: Evolutionarily conserved mediators of immune responses. *Annu. Rev. Immunol. 16*, 225-260.

Grumont, R. J., and Gerondakis, S. (1990). Murine c-rel transcription is rapidly induced in T-cells and fibroblasts by mitogenic agents and the phorbol ester 12-O-tetradecanoylphorbol-13-acetate. *Cell* Growth Differ *1*, 345-350.

Grumont, R. J., Richardson, I. B., Gaff, C., and Gerondakis, S. (1993). rel/NF-kappa B nuclear complexes that bind kB sites in the murine c-rel promoter are required for constitutive c-rel transcription in B-cells. *Cell Growth Differ. 4*, 731-743.

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., Levchenko, A., Scott, M. L., and Baltimore, D. (2002). The IkB-NF-kB signaling module: temporal control and selective gene activation. *Science 298*, 1241-1245.

Horwitz, B. H., Scott, M. L., Cherry, S. R., Bronson, R. T., and Baltimore, D. (1997). Failure of lymphopoiesis after adoptive transfer of NF-kappaB-deficient fetal liver cells. *Immunity 6*, 765-772.

Karin, M., and Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. *Annu. Rev. Immunol. 18*, 621-663.

Karin, M., Cao, Y., Greten, F. R., and Li, Z. W. (2002). NF-kappaB in cancer: from innocent bystander to major culprit. *Nat. Rev. Cancer 2*, 301-310.

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

Kunsch, C., Ruben, S. M., and Rosen, C. A. (1992). Selection of optimal 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.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature 409*, 860-921.

Lee, G. R., Fields, P. E., Griffin, T. J., and Flavell, R. A. (2003). Regulation of the Th2 cytokine locus by a locus control region. *Immunity 19*, 145-153.

Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., Hannett, N. M., Harbison, C. T., Thompson, C. M., Simon, I., et al. (2002). Transcriptional regulatory networks in Saccharomyces cerevisiae. *Science 298*, 799-804.

Lefstin, J. A., and Yamamoto, K. R. (1998). Allosteric effects of DNA on transcriptional regulators. *Nature 392*, 885-888.

Li, Z., Van Calcar, S., Qu, C., Cavenee, W. K., Zhang, M. Q., and Ren, B. (2003). A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells. *Proc. Natl. Acad. Sci. USA 100*, 8164-8169.

Libermann, T. A., and Baltimore, D. (1990). Activation of interleukin-6 gene expression through the NF-kappa B transcription factor. *Mol. Cell. Biol. 10*, 2327-2334.

Lin, Y. S., Carey, M., Ptashne, M., and Green, M. R. (1990). How different eukaryotic transcriptional activators can cooperate promiscuously. *Nature 345*, 359-361.

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

Lois, C., Hong, E. J., Pease, S., Brown, E. J., and Baltimore, D. (2002). Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science 295*, 868-872.

Lomvardas, S., and Thanos, D. (2002). Modifying gene expression programs by altering core promoter chromatin architecture. *Cell 110*, 261-271.

Mai S., a. M. I. L. (1995). The c-myc Protein Represses the Lambda5 and Tdt Initiators. *Nucleic Acids Research 23*, 1-9.

Mechta-Grigoriou, F., Gerald, D., and Yaniv, M. (2001). The mammalian Jun proteins: redundancy and specificity. *Oncogene 20*, 2378-2389.

Meffert, M. K., Chang, J. M., Wiltgen, B. J., Fanselow, M. S., and Baltimore, D. (2003). NF-kB Functions in Synaptic Signaling and Behavior. *Nature Neuroscience*.

Merika, M., Williams, A. J., Chen, G., Collins, T., and Thanos, D. (1998). Recruitment of CBP/p300 by the IFN beta enhanceosome is required for synergistic activation of transcription. *Mol. Cell 1*, 277-287.

Morgenstern, J. P., and Land, H. (1990). A series of mammalian expression vectors and characterisation of their expression of a reporter gene in stably and transiently transfected cells. *Nucleic Acids Research 18*, 1068.

Muller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L., and Harrison, S. C. (1995). Structure of the NF-kappa B p50 homodimer bound to DNA. *Nature 373*, 311-317.

Nissen, R. M., and Yamamoto, K. R. (2000). The glucocorticoid receptor inhibits NFkappaB by interfering with serine-2 phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev. 14*, 2314-2329.

Ohmori, Y., and Hamilton, T. A. (1993). Cooperative interaction between interferon (IFN) stimulus response element and kappa B sequence motifs controls IFN gamma- and

lipopolysaccharide-stimulated transcription from the murine IP-10 promoter. *J. Biol. Chem. 268*, 6677-6688.

Ohmori, Y., and Hamilton, T. A. (1995). The interferon-stimulated response element and a kappa B site mediate synergistic induction of murine IP-10 gene transcription by IFN-gamma and TNF-alpha. *J. Immunol. 154*, 5235-5244.

Pahl, H. L. (1999). Activators and target genes of Rel/NF-kB transcription factors. *Oncogene 18*, 6853-6866.

Perkins, N. D., Agranoff, A. B., Pascal, E., and Nabel, G. J. (1994). An interaction between the DNA-binding domains of RelA(p65) and Sp1 mediates human immunodeficiency virus gene activation. *Mol. Cell. Biol. 14*, 6570-6583.

Perkins, N. D., Edwards, N. L., Duckett, C. S., Agranoff, A. B., Schmid, R. M., and Nabel, G. J. (1993). A cooperative interaction between NF-kappa B and Sp1 is required for HIV-1 enhancer activation. *Embo J. 12*, 3551-3558.

Perkins, N. D., Felzien, L. K., Betts, J. C., Leung, K., Beach, D. H., and Nabel, G. J. (1997). Regulation of NF-kappaB by cyclin-dependent kinases associated with the p300 coactivator. *Science 275*, 523-527.

Phillips, K., and Luisi, B. (2000). The virtuoso of versatility: POU proteins that flex to fit. *J. Mol. Biol. 302*, 1023-1039.

Pierer, M., Rethage, J., Seibl, R., Lauener, R., Brentano, F., Wagner, U., Hantzschel, H., Michel, B. A., Gay, R. E., Gay, S., and Kyburz, D. (2004). Chemokine secretion of rheumatoid arthritis synovial fibroblasts stimulated by Toll-like receptor 2 ligands. *J. Immunol. 172*, 1256-1265.

Ping, D., Boekhoudt, G. H., Rogers, E. M., and Boss, J. M. (1999). Nuclear factor-kappa B p65 mediates the assembly and activation of the TNF-responsive element of the murine monocyte chemoattractant-1 gene. *J. Immunol. 162*, 727-734.

Pomerantz, J. L., and Baltimore, D. (1999). NF-kappaB activation by a signaling complex containing TRAF2, TANK and TBK1, a novel IKK-related kinase. *Embo J. 18*, 6694-6704.

Qin, X. F., An, D. S., Chen, I. S., and Baltimore, D. (2003). Inhibiting HIV-1 infection in human T cells by lentiviral-mediated delivery of small interfering RNA against CCR5. *Proc. Natl. Acad. Sci. USA 100*, 183-188.

Ramji, D. P., and Foka, P. (2002). CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem. J. 365*, 561-575.

Rastinejad, F. (2001). Retinoid X receptor and its partners in the nuclear receptor family. *Curr. Opin. Struct. Biol. 11*, 33-38.

Re, F., and Strominger, J. L. (2001). Toll-like receptor 2 (TLR2) and TLR4 differentially activate human dendritic cells. *J. Biol. Chem. 276*, 37692-37699.

Rogatsky, I., Luecke, H. F., Leitman, D. C., and Yamamoto, K. R. (2002). Alternate surfaces of transcriptional coregulator GRIP1 function in different glucocorticoid receptor activation and repression contexts. *Proc. Natl. Acad. Sci. USA 99*, 16701-16706.

Rogatsky, I., Wang, J. C., Derynck, M. K., Nonaka, D. F., Khodabakhsh, D. B., Haqq, C. M., Darimont, B. D., Garabedian, M. J., and Yamamoto, K. R. (2003). Target-specific utilization of transcriptional regulatory surfaces by the glucocorticoid receptor. *Proc. Natl. Acad. Sci. USA 100*, 13845-13850.

Rothenberg, E. V., and Ward, S. B. (1996). A dynamic assembly of diverse transcription factors integrates activation and cell-type information for interleukin 2 gene regulation. *Proc. Natl. Acad. Sci. USA 93*, 9358-9365.

Saccani, S., Pantano, S., and Natoli, G. (2001). Two waves of nuclear factor kappaB recruitment to target promoters. *J. Exp. Med*. *193*, 1351-1359.

Saccani, S., Pantano, S., and Natoli, G. (2003). Modulation of NF-kappaB activity by exchange of dimers. *Mol. Cell 11*, 1563-1574.

Sakaguchi, S., Negishi, H., Asagiri, M., Nakajima, C., Mizutani, T., Takaoka, A., Honda, K., and Taniguchi, T. (2003). Essential role of IRF-3 in lipopolysaccharide-induced interferon-beta gene expression and endotoxin shock. *Biochem. Biophys. Res. Commun. 306*, 860-866.

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. USA 97*, 12705-12710.

Scott, M. L., Fujita, T., Liou, H. C., Nolan, G. P., and Baltimore, D. (1993). The p65 subunit of NF-kappa B regulates I kappa B by two distinct mechanisms. *Genes Dev. 7*, 1266-1276.

Serbina, N. V., Kuziel, W., Flavell, R., Akira, S., Rollins, B., and Pamer, E. G. (2003). Sequential MyD88-independent and -dependent activation of innate immune responses to intracellular bacterial infection. *Immunity 19*, 891-901.

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.

Sheppard, K. A., Rose, D. W., Haque, Z. K., Kurokawa, R., McInerney, E., Westin, S., Thanos, D., Rosenfeld, M. G., Glass, C. K., and Collins, T. (1999). Transcriptional activation by NF-kappaB requires multiple coactivators. *Mol. Cell. Biol. 19*, 6367-6378.

Takeda, K., Kaisho, T., and Akira, S. (2003). Toll-like receptors. Annu Rev Immunol *21*, 335-376.

Thanos, D., and Maniatis, T. (1995a). Identification of the rel family members required for virus induction of the human beta interferon gene. *Mol. Cell. Biol. 15*, 152-164.

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

Trimarchi, J. M., and Lees, J. A. (2002). Sibling rivalry in the E2F family. *Nat. Rev. Mol. Cell. Biol. 3*, 11-20.

Weinmann, A. S., Yan, P. S., Oberley, M. J., Huang, T. H., and Farnham, P. J. (2002). Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes Dev. 16*, 235-244.

Wietek, C., Miggin, S. M., Jefferies, C. A., and O'Neill, L. A. (2003). Interferon regulatory factor-3-mediated activation of the interferon-sensitive response element by Toll-like receptor (TLR) 4 but not TLR3 requires the p65 subunit of NF-kappa. *J. Biol. Chem. 278*, 50923-50931.

Wolfe, S. A., Nekludova, L., and Pabo, C. O. (2000). DNA recognition by Cys2His2 zinc finger proteins. *Annu. Rev. Biophys. Biomol. Struct. 29*, 183-212.

Xu, Y., Davidson, L., Alt, F. W., and Baltimore, D. (1996). Deletion of the Ig kappa light chain intronic enhancer/matrix attachment region impairs but does not abolish V kappa J kappa rearrangement. *Immunity 4*, 377-385.

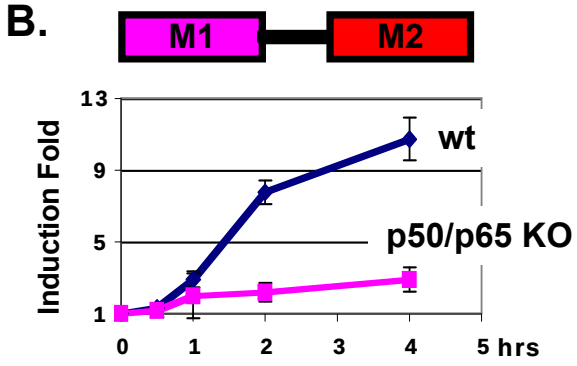
Yamamoto, M., Sato, S., Hemmi, H., Hoshino, K., Kaisho, T., Sanjo, H., Takeuchi, O., Sugiyama, M., Okabe, M., Takeda, K., and Akira, S. (2003). Role of adaptor TRIF in the MyD88-independent toll-like receptor signaling pathway. *Science 301*, 640-643.

Yamamoto, M., Sato, S., Hemmi, H., Sanjo, H., Uematsu, S., Kaisho, T., Hoshino, K., Takeuchi, O., Kobayashi, M., Fujita, T., et al. (2002). Essential role for TIRAP in activation of the signalling cascade shared by TLR2 and TLR4. *Nature 420*, 324-329.

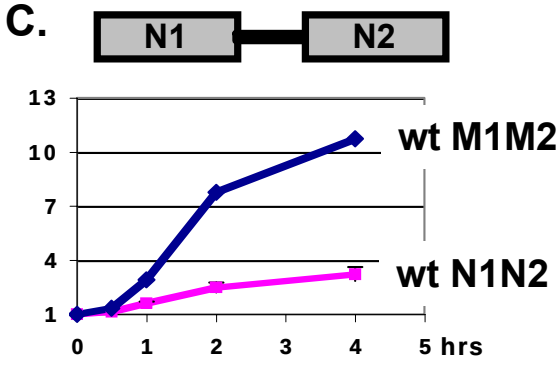
A.

Gene	Mouse	Human
MCP-1	GGGAACTTCC GGGAATTTCC	GGGAACTTCC GGGAATTTCC
IkBa	GGAAATTCCC GGGGAAGTCC	GGAAATTCCC GGGGAAGTCC
Fas	GGGAATGCCC	GGGAATGCCC
MIP2	GGGAATTTCC	GGGAATTTCC
M-CSF	GGAAAGTCCC	GGAAAGTCCC
IP-10	GGGAAATTCC GGGACTTCCC	GGGAAATTCC GGGACTTCCC
RANTES	GGGAGTTTCC	GGGAGTTTCC
LIF	GGGGATCCCG	GGGGATCCCG
IFN§	GGGAAATTCC	GGGAAATTCC
IL-6	GGATTTTCCC	GGATTTTCCC
JunB	GGGGCTTTCC	GGGGCTTTCC

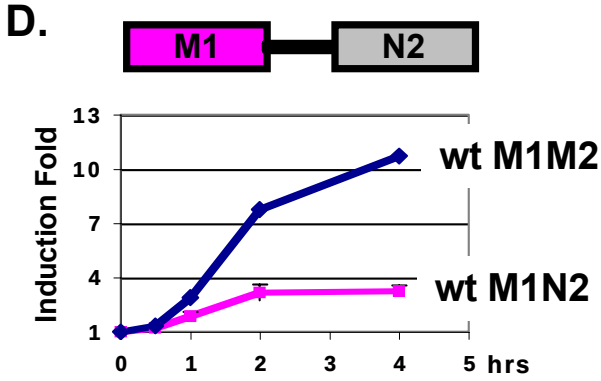
B.

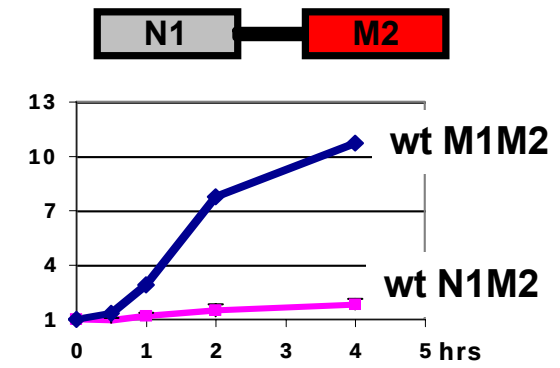


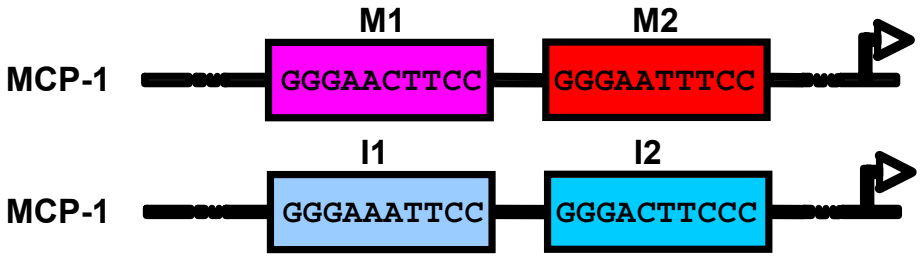
C.

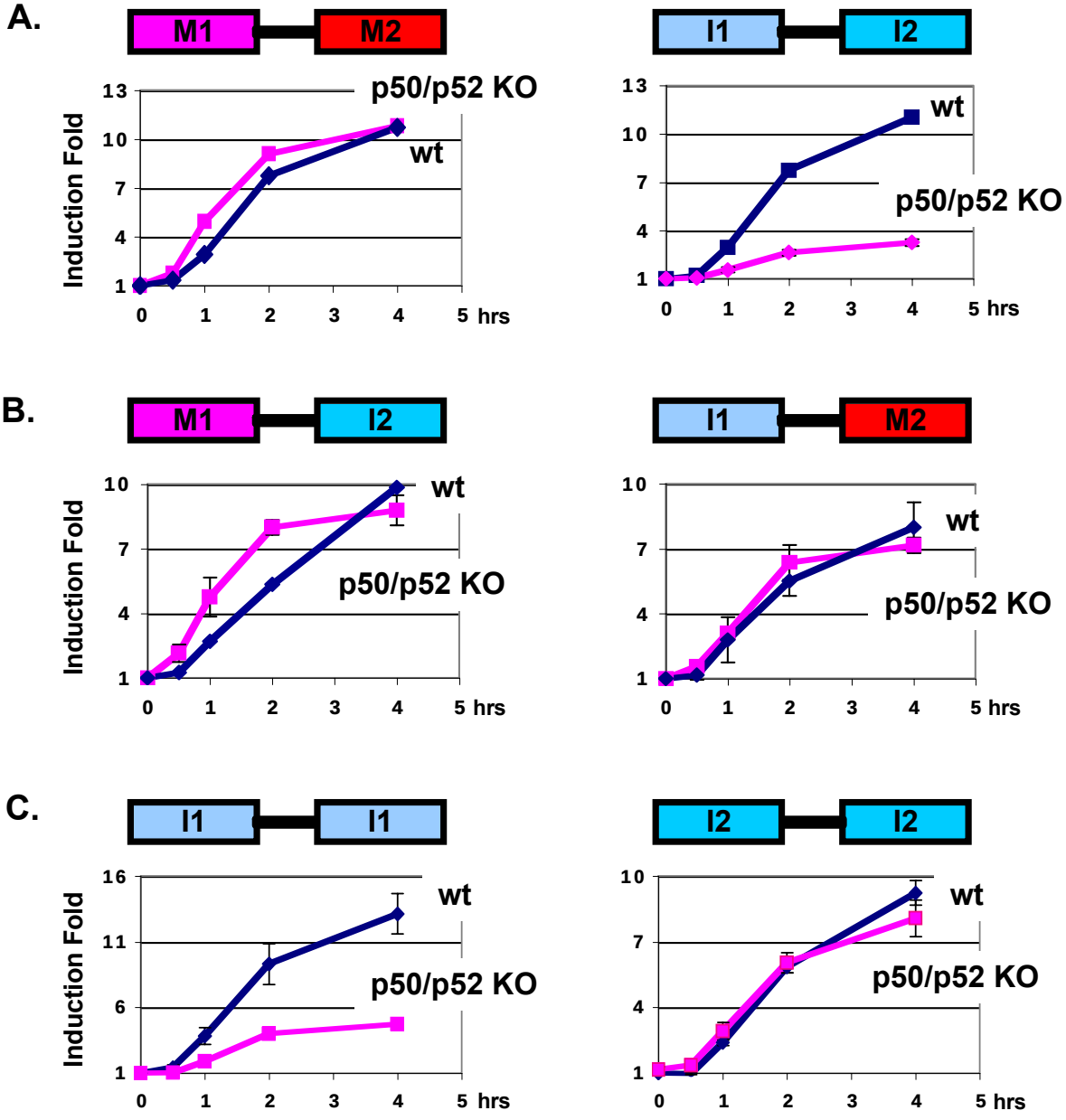


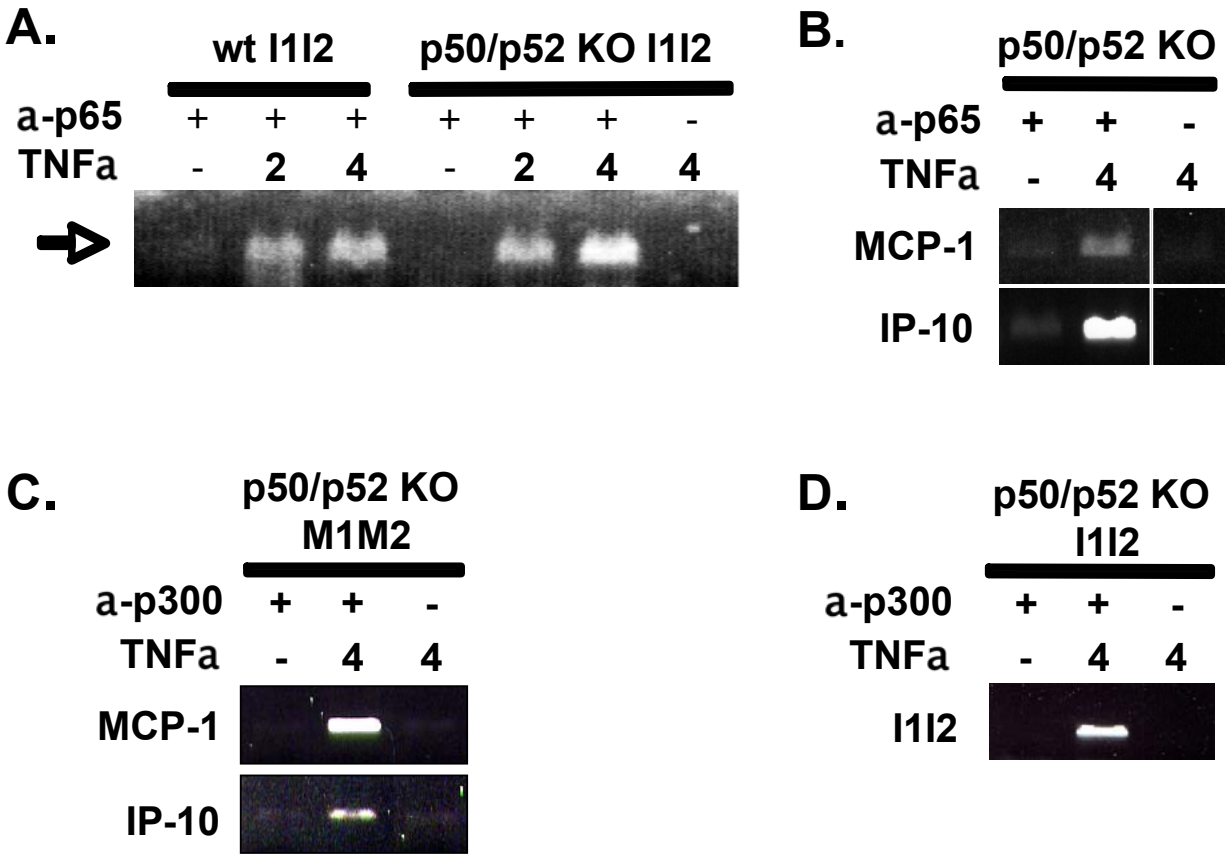
D.



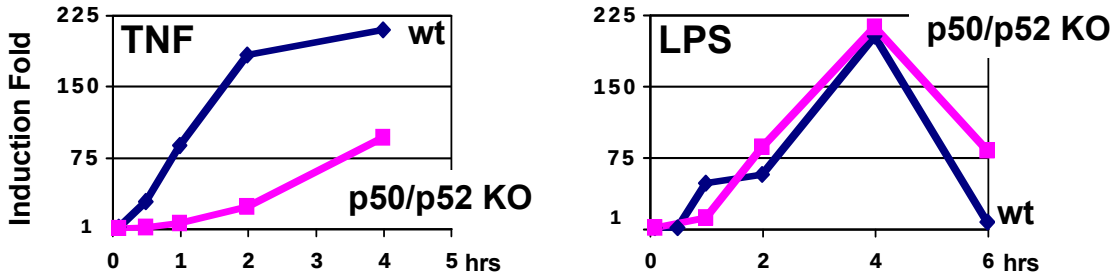



M1
M2
MCP-1
GGGAACTTCC
GGGAATTTCC
I1
I2
MCP-1
GGGAAATTCC
GGGACTTCCC

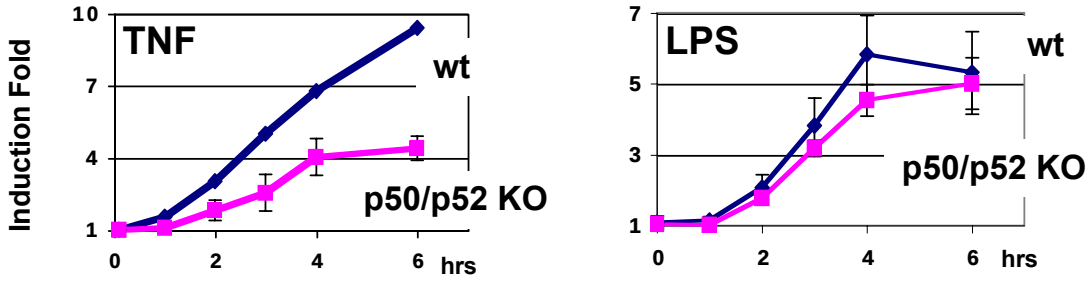

A.
M1
M2
Induction Fold
p50/p52 KO
wt
5 hrs
I1
I2
wt
p50/p52 KO
5 hrs
B.
M1
I2
Induction Fold
wt
p50/p52 KO
5 hrs
I1
M2
wt
p50/p52 KO
5 hrs
C.
I1
I1
Induction Fold
wt
p50/p52 KO
5 hrs
I2
I2
wt
p50/p52 KO
5 hrs


A.
wt I1I2
p50/p52 KO I1I2
a-p65
+ + + + + + -
TNFa
- 2 4 - 2 4 4
B.
p50/p52 KO
a-p65
+ + -
TNFa
- 4 4
MCP-1
IP-10
C.
p50/p52 KO
M1M2
a-p300
+ + -
TNFa
- 4 4
MCP-1
IP-10
D.
p50/p52 KO
I1I2
a-p300
+ + -
TNFa
- 4 4
I1I2

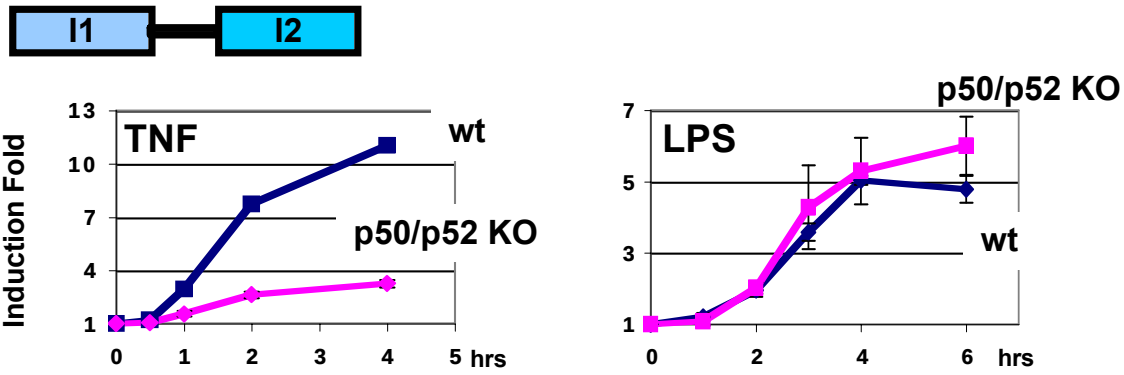
A. Endogenous IP-10 Transcript

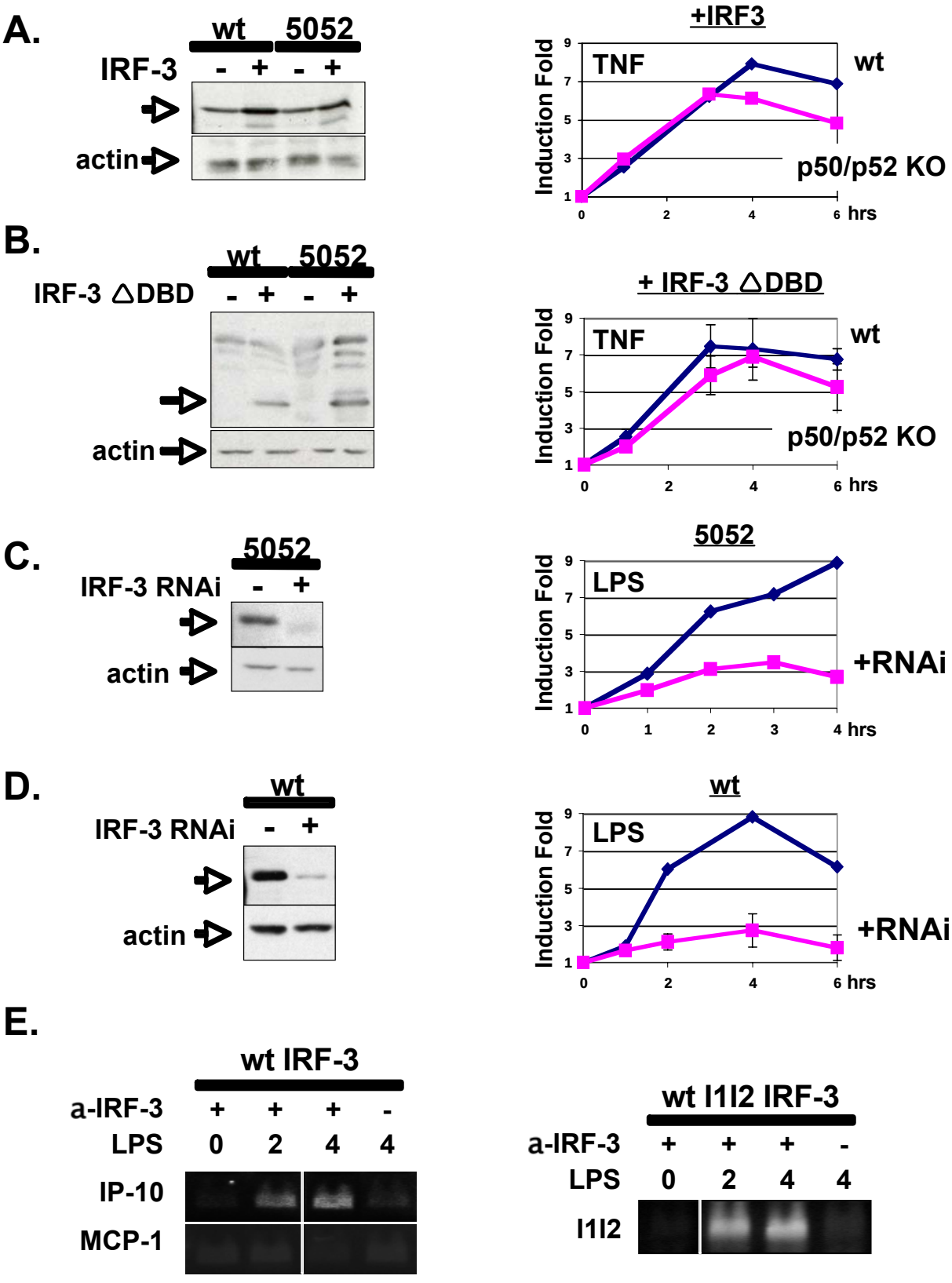


B. IP-10 Transgene

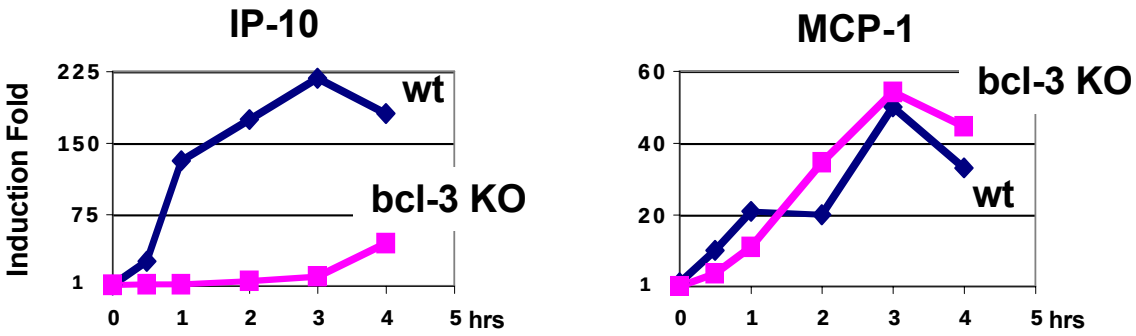


C.

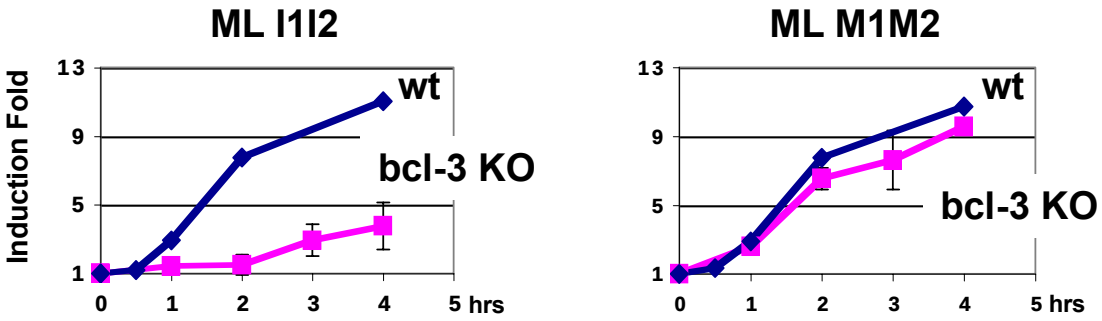


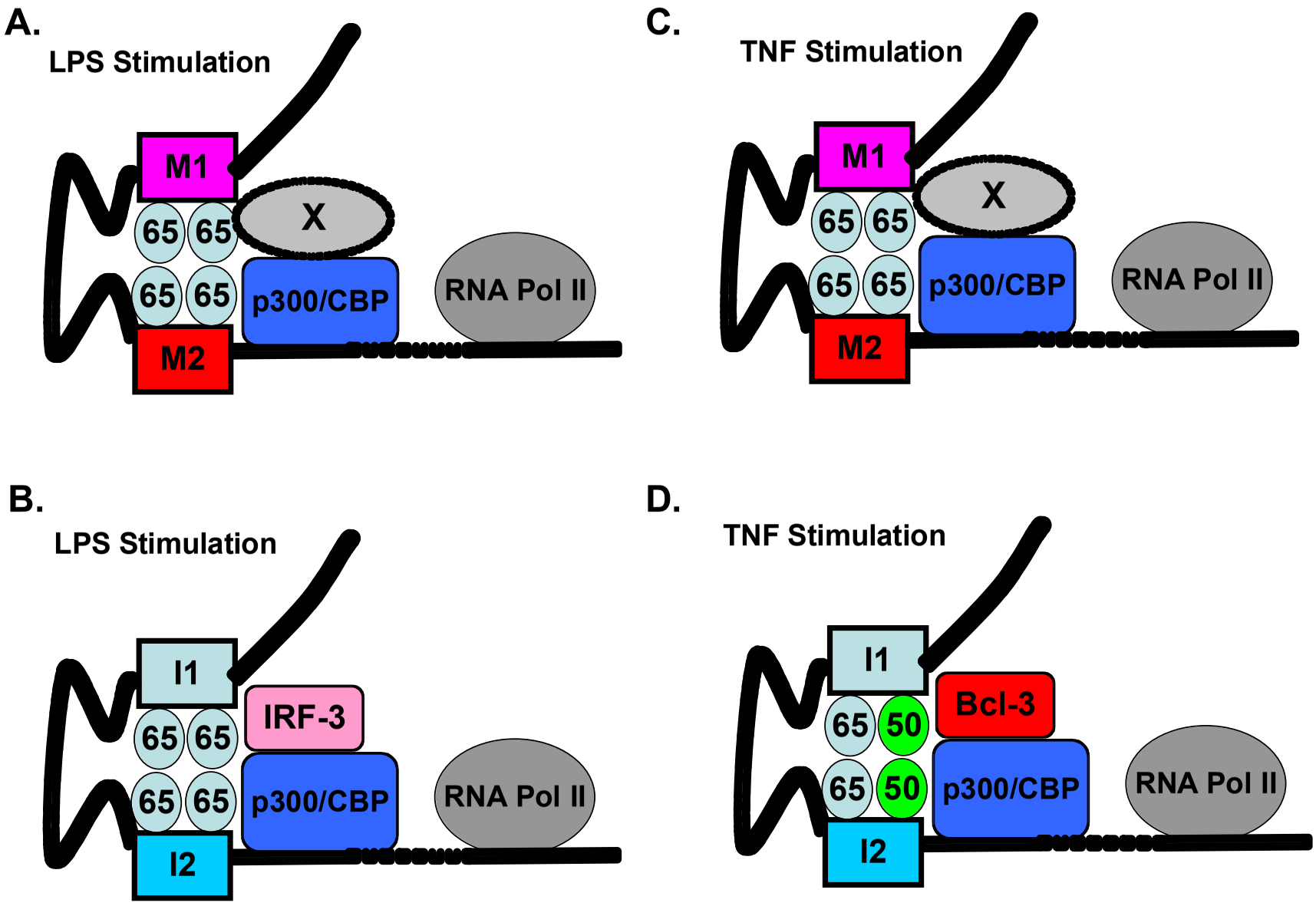

A.
wt
5052
IRF-3
- + - +
actin
+IRF3
TNF
wt
p50/p52 KO
Induction Fold
hrs
B.
wt
5052
IRF-3 ΔDBD
- + - +
actin
+ IRF-3 ΔDBD
TNF
wt
p50/p52 KO
Induction Fold
hrs
C.
5052
IRF-3 RNAi
- +
actin
5052
LPS
+RNAi
Induction Fold
hrs
D.
wt
IRF-3 RNAi
- +
actin
wt
LPS
+RNAi
Induction Fold
hrs
E.
wt IRF-3
a-IRF-3 + + + -
LPS 0 2 4 4
IP-10
MCP-1
wt I1I2 IRF-3
a-IRF-3 + + + -
LPS 0 2 4 4
I1I2

A.


IP-10
Induction Fold
225
150
75
1
wt
bcl-3 KO
0
1
2
3
4
5 hrs
MCP-1
60
40
20
1
bcl-3 KO
wt
0
1
2
3
4
5 hrs

B.


ML I1I2
Induction Fold
13
9
5
1
wt
bcl-3 KO
0
1
2
3
4
5 hrs
ML M1M2
13
9
5
1
wt
bcl-3 KO
0
1
2
3
4
5 hrs


A.
LPS Stimulation
M1
65
65
X
65
65
p300/CBP
RNA Pol II
M2
C.
TNF Stimulation
M1
65
65
X
65
65
p300/CBP
RNA Pol II
M2
B.
LPS Stimulation
I1
65
65
IRF-3
65
65
p300/CBP
RNA Pol II
I2
D.
TNF Stimulation
I1
65
50
Bcl-3
65
50
p300/CBP
RNA Pol II
I2

CHAPTER 5: FUTURE STUDIES

There are two major directions for future work in the lab. 1. To understand how universal is this mechanism of regulation? 2. To extend our studies and use biochemical and biophysical methods to understand mechanistically how the κB site sequence determines co-activator requirements.

Aim 1. Global mechanism studies

We would like to understand if κB family member and co-activator requirements generated by the κB site sequence remain the same across different cell types and promoters. We postulate that the κB site may place bound κB dimer(s) in a particular conformation that interacts most favorable with other important control elements for a particular promoter. Evolution would ensure that preferred sequences are promoted and conserved. If this hypothesis holds true, the sequence of the κB site may predict potential interaction partners between NF-κB and other co-activators/transcription factors. For example, the κB site sequences found in the IP-10 promoter may position bound κB dimers to interact favorably with IRF-3 or Bcl-3. We could then catalog potential interaction partners for each κB site sequence. Ultimately, this information would help decipher the promoter code and gene regulatory networks.

Some questions to address are:

1. Are gene promoter requirements for specific κB family members and co-activators the same across different cell types?

2. Our c-fos minimum promoter experiments suggested that κB family members requirements are intrinsic to the κB site sequence and cryptic sites were not responsible for this behavior. Do genes

that have the exact same κB site sequences respond in similar fashions? Our IP-10 κB site sequences are an ideal starting point for a genome analysis. The presence of other transcription factor binding sites on the "identified" genes may or may not affect gene regulation. Either way, it will be interesting to see if different patterns or consistencies among genes with identical κB binding site sequences can be formed.

3. Do other dual κB-binding site containing genes behave in a similar way? Our system is easily amendable to the study of other promoters. The more examples we can test, the better we will understand what is going on.

4. Do single κB-binding site containing promoters behave in a similar manner? A single site may reduce the complexity of the promoter code and allow for easier dissection of potential.

5. Do other transcription factor families behave in a similar way? As mentioned above, outside the NHR family, little is know about how other eukaryotic transcription factor families determine specificity. Other families are similarly set-up to NF-κB (i.e. many family members recognize the same binding sites), and this mechanism may answer why marked DNA-binding specificity rules have not been identified.

Aim 2: Biochemical and Biophysical Studies

1. What is factor X? Previously, we showed that MCP-1 κB sites are capable of promoting gene expression without Bcl-3 or IRF-3 co-activators and the presence of either MCP-1 κB site is

capable of bypassing the Bcl-3 requirement. In our model, we proposed that each κB site serves a specific non-redundant function and suggest that either MCP-1 κB site is capable of recruiting factor "X", a co-activator needed for gene activation.

2. Characterization of the interactions between p50:p65 heterodimers and its new co-activators, Bcl-3 and IRF-3. While our results have established their interactions as a physiologically relevant control mechanism, we lack detailed information regarding their interactions, which may or may not involve another cofactor or meditating protein.

3. Biophysical studies to address whether the conformation of bound NF-κB dimers directly determines co-activator requirements via Biacore technology and crystal structures.
Significant conformational differences are expected between p50:p65 heterodimers bound to M1M2 and I1I2 that will prove that the sequence of the κB site affects the conformation of the bound dimer. If the structure of p50:p65 heterodimers bound to M1M2 and I1I2 are solved, the analysis should be extended, and an attempt to solve a structure with p50:p65 heterodimers bound to M1M2 or I1I2 and their appropriate co-activators will be made.

Concluding Remarks

The field of gene regulation has had a seemingly natural progression. The field started with bacteria and has now progressed to study higher eukaryotes. By steadily increasing genome sizes, some of the original concepts have also increased in complexity. In bacteria, we began with two molecules whose binding sites were close to the transcription start site. In larger genomes, we discovered that binding sites could be moved farther away. Finally in higher eukaryotes, the model

promoter of gene regulation has (at least) four different transcription factor families interacting for gene activation to occur. My work has built on this foundation and addressed a long-standing question in the field, how do individual members in a particular transcription factor family function on a given gene promoter.

Other concepts of gene regulation have remained the same but presented in a different flavor. In bacteria, allostery between allolactase and repressor controlled gene activation. My work identified a situation where DNA may act as an allosteric regulator of NF-κB to determine gene activation. Instead of a protein-protein allosteric interaction, we are characterizing a DNA-protein allosteric interaction. I'm still a ways away from proving this to be empirically true, but as Monod said: "There is in science, however, quite a gap between belief and certainty. But would one ever have patience to wait and to establish the certainty if the inner conviction were not already there?"

While I am trying to establish certainty, I still have much to learn from bacterial gene regulation. Two recent findings, one for *lacZ* and one for lambda repressor, illustrates this best and may have important implications for the control of NF-κB. As described above, crystal structures recently demonstrated that the sequence of DNA determines the conformation of bound *lac*-Repressor and affects *lac*-Repressor's function. Depending on the sequence of DNA, the DNA-binding domain of repressor can actually adopt two strikingly different conformations. We postulate that this phenomenon applies to NF-κB. Also, oligomerization of the lambda repressor was shown to be a significant and biological relevant mechanism of regulation for the lambda switch. As described above, we find this to be an attractive hypothesis for regulation in the NF-κB family as well.

All in all, this is an exciting time to be in the field of gene regulation.