Topic Introduction

Identifying *cis*-Acting DNA Elements within a Control Region

Michael F. Carey, Craig L. Peterson, and Stephen T. Smale

Computational methods can be used to identify DNA sequence motifs that have been conserved through evolution, as well as motifs that correspond to recognition sites for known DNA-binding proteins. These computational methods, when combined with chromatin immunoprecipitation and other basic experiments, can provide preliminary insight into the elements and factors that regulate a gene of interest. When pursuing a more complete understanding of a control region of interest, a comprehensive mutant analysis should generally be performed as a critical step toward more advanced functional studies. This article describes strategies for such a comprehensive analysis. It also summarizes the insights provided by a comprehensive mutant analysis versus a phylogenetic analysis.

INTRODUCTION AND OVERVIEW

Strategies and methods are available for identifying transcriptional control regions, including promoters and a variety of distant control regions. To begin to elucidate the mechanisms by which a control region regulates transcription, the relevant *cis*-acting DNA sequence elements must next be delineated and the *trans*-acting protein factors that interact with those elements defined. In the past, we strongly recommended that researchers identify important DNA elements in a control region of interest before attempting to identify relevant transcription factors. We further recommended a comprehensive mutant analysis as an essential first step toward identifying important DNA sequence elements within the control region. At that time, a reliable technique for monitoring the binding of specific proteins to a control region of interest in living cells had not been established. Therefore, studies of protein binding were generally performed *in vitro*; a considerable amount of work was required to determine whether the same protein bound the control region *in vivo* and if it was important for the control region's function in a physiological setting. Comprehensive mutant analyses were also recommended because computational strategies for identifying DNA motifs that might contribute to the regulation of a gene (based on evolutionary conservation or the presence of a recognition site for a known transcription factor) were primitive.

There have, however, been dramatic advances in chromatin immunoprecipitation (ChIP) technology for monitoring protein–DNA interactions *in vivo* and even greater advances in phylogenetic analyses of DNA sequences. Therefore, we now recommend that researchers begin to dissect a control region of interest by performing a phylogenetic analysis to identify DNA sequence motifs that have been strongly conserved through evolution. Because complete or nearly complete genome sequences are available for such a large number of organisms, strong conservation of a motif (through all of vertebrate evolution, for example) provides much stronger evidence of functional relevance than could be provided previously, when only two or three species could be compared. After motifs that are well-conserved through evolution are identified, each motif can be analyzed computationally (see

Adapted from *Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques*, 2nd edition, by Michael F. Carey, Craig L. Peterson, and Stephen T. Smale. CSHL Press, Cold Spring Harbor, NY, USA, 2009.

© 2012 Cold Spring Harbor Laboratory Press

Cite this article as *Cold Spring Harbor Protoc*; 2012; doi:10.1101/pdb.top068171

below) to determine whether it corresponds to a consensus or near-consensus binding site for a known protein or family of proteins. If antibodies are available for a candidate protein, ChIP experiments can be performed to determine whether the candidate associates with the control region in vivo. If positive results are obtained, functional studies can be initiated to determine whether the protein contributes to the activity of the control region. Functional studies generally include transcription factor loss-of-function or gain-of-function experiments or experiments in which the predicted DNA-binding site for the protein is disrupted.

Although this strategy can provide considerable insight into the transcription factors that regulate a gene of interest, there are a number of notable caveats, which lead us to recommend it with some hesitation. First, the initial phylogenetic analysis will identify only those DNA elements that have been well conserved through evolution. Species-specific DNA elements, which must exist in some control regions, will be missed in a phylogenetic analysis. Unfortunately, the extents to which regulatory strategies and critical DNA elements for a given gene vary from species to species remain unknown. The predictions of researchers vary widely regarding the percentage of key regulatory elements that will be missed if evolutionary conservation is used as the sole criterion for identifying important elements. Moreover, many DNA-binding proteins can accommodate substantial changes in a DNA recognition sequence without reducing binding affinity, which further increases the possibility that at least some critical elements will be missed in a phylogenetic analysis.

A related and even greater concern is that computational programs for predicting which known DNA-binding proteins interact with a DNA motif remain primitive. The quality of these programs depends largely on the quality of the experimental data used to generate them. One obvious problem is that consensus recognition sites for only a small number of DNA-binding proteins within the proteome have been defined experimentally. Furthermore, comprehensive information about the range of sequences that can be recognized by a DNA-binding protein with sufficient affinity for functional activity is not available for any eukaryotic DNA-binding protein. Related to this problem, some proteins function through low-affinity sites that diverge considerably from the consensus recognition sequence. These proteins generally bind cooperatively with other DNA-binding proteins. A final problem is that many DNA-binding proteins are members of large multiprotein families, with multiple family members recognizing similar or identical DNA sequences. Thus, even if the computational program successfully predicts which protein family is capable of binding an important evolutionarily conserved DNA motif, it may be difficult to predict which family member (or members) is responsible for the function of the motif in its native genomic context. Because of these and other limitations for predicting which proteins interact with an evolutionarily conserved DNA motif, it is quite possible that the relevant DNA-binding protein will not be revealed by a computational approach. In these instances, it may be necessary to use one of the methods outlined in *Experimental Strategies for the Identification of DNA-Binding Proteins* (Carey et al. 2012a) or *Experimental Strategies for Cloning or Identifying Genes Encoding DNA-Binding Proteins* (Carey et al. 2012b), along with other strategies for assessing functional relevance.

The next step in our recommended strategy is to use ChIP experiments to examine the ability of candidate proteins to bind the control region of interest in vivo. It is important to keep in mind that ChIP is not a functional assay. Therefore, evidence that a protein binds to a control region in vivo provides limited insight into the functional relevance of the interaction. ChIP is further limited by its dependence on high-quality antibodies for the protein of interest and on protein–DNA interactions that are amenable to ChIP analysis. Some proteins are difficult to study using ChIP assays for a variety of reasons, including transient binding or inefficient cross-linking to DNA or epitope masking.

Despite these serious caveats, phylogenetic analyses and ChIP experiments are now recommended as initial steps toward functional studies because they can provide useful knowledge rapidly. If the goal is to identify a subset of regulatory elements and DNA-binding proteins for a control region of interest, and if the protein candidates suggested by the computational approaches are of sufficient interest, the researcher may wish to focus future studies on these elements and proteins, without performing a more comprehensive analysis of the control region. On the other hand, if the researcher is committed to understanding fully the regulation of a gene, a comprehensive mutant analysis to identify all

elements that contribute to the function of a control element in an appropriate functional assay (e.g., transient transfection with promoter–reporter plasmids) is strongly recommended. This analysis will provide insight into which of the evolutionarily conserved elements are critical for the control region’s activity in the chosen assay and whether critical DNA elements might have been missed in the phylogenetic analysis, owing to insufficient conservation through evolution. However, as described below, comprehensive mutant analyses have their own significant limitations.

We do not recommend that researchers simply use biochemical methods to identify proteins that bind *in vitro* to a control region of interest without first pinpointing the locations of individual DNA motifs that are likely to be functionally important, through either phylogenetic analysis or mutagenesis. It is well known that nuclear extracts contain a large number of proteins that are capable of binding to any DNA fragment of 100–400 bp (the length of a typical promoter or enhancer). Therefore, efforts to find proteins capable of binding a DNA fragment of interest will undoubtedly succeed, but the probability of preferentially identifying relevant DNA-binding proteins using this approach is very low.

It is important to keep in mind that regardless of whether a researcher begins with the phylogenetic analysis leading to ChIP experiments or with a comprehensive mutant analysis, the results obtained will provide only an initial view of the mechanisms regulating the gene. The experimental strategies described in **Experimental Strategies for the Identification of DNA-Binding Proteins** (Carey et al. 2012a) and **Experimental Strategies for Cloning or Identifying Genes Encoding DNA-Binding Proteins** (Carey et al. 2012b) may be required to identify proteins capable of binding the DNA motifs that are identified. Furthermore, for virtually all studies, it will be necessary to carefully evaluate the functional relevance of each protein–DNA interaction.

IDENTIFICATION OF CONTROL ELEMENTS BY COMPREHENSIVE MUTANT ANALYSIS

As described above, we currently recommend that researchers begin with a phylogenetic analysis, computational analysis of recognition motifs for known DNA-binding proteins, and possibly ChIP experiments to gain initial knowledge about the elements and factors responsible for the function of a control region of interest. In this article, we will focus on preferred strategies for a comprehensive mutant analysis. This discussion has been included because, when researchers new to the transcription field attempt to perform this kind of analysis, the answers to many detailed technical questions often are not obvious. We also note that comprehensive mutant analyses are being performed less frequently, as researchers shift their emphasis to phylogenetic analyses, ChIP assays, and transcription factor loss-of-function studies. Nevertheless, because each of these approaches has substantial limitations, detailed mutant analyses will eventually be required to achieve an advanced understanding of the mechanisms regulating any eukaryotic gene.

It is important to emphasize that the potential benefits of a comprehensive mutant analysis must be balanced against three limitations. First, a meaningful analysis requires substantially more effort than phylogenetic or ChIP analyses. Fortunately, the time required for a comprehensive mutant analysis is decreasing as subcloning and *in vitro* mutagenesis technologies improve. A second and more significant limitation is that the DNA elements identified will only include those that are required for the control region to function in the chosen assay. Transient transfection assays may allow most of the elements required for the function of a control region to be identified. However, control elements may exist within a promoter or enhancer that functions only in the context of a native chromatin environment. Furthermore, the high plasmid copy number that often exists in a cell following transient transfection may cause some important control elements to be overlooked, or less important elements to predominate. A stable transfection assay may have a better chance of revealing all of the important elements within a control region. However, when used for a comprehensive mutant analysis, stable transfection assays are extremely time consuming and also may fail to reveal the functions of all DNA motifs that are required in a native genomic context. A third limitation of the mutant analysis concerns the possible redundancy of DNA elements within a control

region. If redundancies exist, it can be difficult to determine the boundaries of the control region and identify important DNA elements.

The advantages and disadvantages of the comprehensive mutant analysis are difficult to balance. On the one hand, a comprehensive analysis may reveal important regulatory elements missed in a phylogenetic analysis or computational analysis of recognition sites for known proteins. On the other hand, the mutant analysis requires much more effort, and the information obtained is likely to remain incomplete. Therefore, this issue must be evaluated with respect to the overall goals of the study, along with the available resources and time commitment (see above).

STRATEGIES FOR A COMPREHENSIVE ANALYSIS

Most basic mutagenesis studies include a combination of deletion and substitution mutants. As an example, we describe the dissection of the promoter for murine *Il12b*, which encodes the p40 subunit of the heterodimeric cytokines interleukin-12 (IL-12) and interleukin-23 (IL-23) (Plevy et al. 1997). We focus initially on the strategy used to delineate the DNA elements required for *Il12b* promoter activity following induction by an extracellular stimulus; this approach can be applied to any control region, regardless of its mode of regulation. The issue of which elements are directly responsible for induction is discussed later in this article.

Dissecting the *Il12b* promoter began with a deletion analysis, performed primarily to define the boundaries of the functional promoter. The individual control elements within those boundaries were then identified by scanning the region with a series of clustered substitution mutations of 6–10 bp each. Small deletions were not used because they have the potential to alter the alignment of control elements, which may be important for proper regulation. The third step was to analyze each of the control elements with more refined substitution mutations of 3 bp, to determine their boundaries and to determine whether each represents a binding site for one protein or a composite site for two or more proteins. The issues considered when designing each step in this strategy are presented below, along with the information obtained. Although this example involves the dissection of a promoter, a similar strategy can be envisioned for enhancers and other distant control regions.

The *Il12b* gene is expressed in macrophages and dendritic cells that have been activated by bacterial products, such as lipopolysaccharide (LPS). The initial goal of the analysis was to dissect the mechanism of gene induction by LPS. Because most inducible cytokine genes that have been characterized are regulated primarily by promoter sequences, it was considered likely that the *Il12b* promoter would, at least in part, be responsible for induction. Before our analysis, Trinchieri and colleagues used nuclear run-ons to show that induction is regulated primarily at the level of transcription initiation (Ma et al. 1996). The transcription initiation site for the murine gene had been mapped to a location ~25 bp downstream from an AT-rich sequence that was likely to function as a TATA box (Murphy et al. 1995).

A 405-bp promoter fragment (extending from –350 to +55 relative to the start site) was fused to luciferase and chloramphenicol acetyltransferase (CAT) reporter genes in standard reporter vectors (Plevy et al. 1997). CAT reporter genes generally are not recommended because the CAT reporter assay is more time consuming and less sensitive than other reporter assays, such as luciferase. However, the CAT reporter was found to be preferable in this study because the firefly luciferase cDNA available at the time contained sequences that led to nonspecific induction of luciferase activity in LPS-stimulated macrophages (Plevy et al. 1997). For *Il12b* promoter insertion, the upstream boundary of –350 was chosen because most promoters include key elements within a few hundred base pairs of the start site. The downstream boundary of +55 is near the translation initiation codon and was chosen because some genes contain important promoter elements in the untranslated leader. Following transfection of these plasmids into the RAW 264.7 macrophage cell line and activation by LPS, the 405-bp fragment was found to be sufficient for strong, inducible promoter activity (Plevy et al. 1997). Therefore, a mutagenesis strategy was needed to identify the important control elements within this fragment.

Deletion Analysis

The first step of the mutant analysis was to generate and analyze a series of 5' deletion mutants (Plevy et al. 1997). Deletions were prepared by polymerase chain reaction (PCR), using a forward primer spanning the desired end point of the deletion and a reverse primer spanning nucleotide +55. Both primers contained sequences that generated restriction sites adjacent to the end points of the promoter fragment, allowing endonuclease cleavage and direct insertion into the reporter vector. Because the mutations were generated by PCR, it was necessary to sequence the final plasmid inserts to ensure that unwanted point mutations were not present.

Deletion mutants were prepared and analyzed before substitution mutants because the deletion analysis defines the minimal sequence that supports full activity. By determining the minimal sequence, the number of substitution mutants that subsequently must be prepared is minimized. For example, if the entire 405-bp fragment were essential for activity, 41 10-bp substitution mutations would be needed to scan the region for important elements. In contrast, if 200 bp at the 5' end of this fragment could be eliminated without a significant effect on activity, only 21 10-bp mutations would be needed to scan the functionally relevant region. Indeed, the results revealed that 100% of the promoter activity was retained with a fragment extending from –200 to +55, and 40% was retained with a fragment extending from –150 to +55 (Fig. 1). Furthermore, promoter activity was not significantly enhanced when sequences extending to –800 were included.

A few points regarding the interpretation of these initial deletion results are noteworthy. First, the data in Figure 1 show that deletion of sequences between –350 and –150 had small effects on promoter activity. For example, a deletion from –250 to –215 reduced activity twofold, a deletion from –215 to –200 enhanced activity twofold, and a deletion from –200 to –150 again reduced activity twofold. The deleted sequences that led to these effects might contain important positive and negative control elements. Alternatively, the twofold differences may be irrelevant to promoter activity. Each deletion results in the fusion of a sequence from the *Il12b* locus to vector sequences. The vector sequences

FIGURE 1. Basic deletion analysis for the murine *Il12b* promoter. (A) Deletions of sequences between –350 to +55 and –150 to +55 and their effect, expressed as % of wild-type promoter activity. (B) Deletions of sequences beginning between –150 and –23 and their effect, expressed as % of wild-type promoter activity. (Adapted, with permission, from Plevy et al. 1997, ©American Society for Microbiology.)

could have modest effects on promoter activity, either positive or negative, when fused to different nucleotides within the promoter. Thus, it cannot simply be assumed that every twofold difference corresponds to a relevant promoter element. To determine whether a bona fide element exists, for example, between –250 and –215, 10-bp substitution mutations scanning this 35-bp region could be introduced into the –350 to +55 promoter fragment. If one or two of these mutations recapitulate the twofold loss in activity, a relevant control element might exist within this region. The putative element could be localized more precisely with smaller substitution mutants, and its mechanism of action could then be analyzed in detail. If the effect is not observed with the substitution mutations, the effect observed in the deletion analysis may be irrelevant.

A second point regarding the deletion data is that the standard deviations determined for some of the mutants are quite large. The reason for the considerable variability is that data were used from multiple experiments performed on different days with different DNA preparations. The standard deviations would have been much lower if they had been derived from multiple independent transfections performed on the same day or with the same DNA preparation. However, the effect of each mutation varies to some degree from one DNA preparation to another and from day to day, possibly because of different concentrations of contaminants in different DNA preparations or differences in the growth state, health, and transfection efficiency of the cells. Although larger standard deviations are obtained when these variations are documented, they lead to a more accurate presentation of the data. If the standard deviations had been derived from experiments performed with only one DNA preparation or on only one day, the data could have been less accurate and perhaps misleading.

Finally, the data for each mutant are presented as a percentage of the wild-type promoter activity following induction (i.e., the activity of the induced wild-type promoter is set as 100% with the activity of each mutant following induction determined relative to wild type), not as the fold-activation by LPS. This latter number would rely on the validity of the uninduced signal. Because the uninduced signals were quite close to background, very little confidence can be placed in their validity. Further comments regarding the documentation of inducibility are included below.

The more relevant deletions appeared to be those that extend past –150. As indicated in the graph (Fig. 1), deletion from –143 to –135 reduced promoter activity to 20% of wild type (i.e., the –350 to +55 fragment), deletion from –126 to –123 reduced activity to 10%, and deletion from –102 to –97 reduced activity to near background. These results suggest that important elements are likely to exist between –143 and –97. Furthermore, equally important elements may exist between –97 and +55. Because activity was reduced to background levels following deletion to –97, these data provide no information about the existence of important elements downstream from –97. Substitution mutants or 3' deletion mutants are needed to determine whether important control elements exist in this downstream region.

For the *Il12b* study, deletion mutants were not prepared from the 3' end of the promoter (i.e., sequentially deleting sequences from +55 toward the transcription start site). This is because the –150 to +55 sequence is of reasonable size to dissect by substitution mutant analysis. Furthermore, because only 55 bp of untranslated leader was included, only a small number of nucleotides could be deleted from the 3' end without affecting the core promoter elements, including the TATA box and potential start-site sequence. Because deletion of these sequences could influence the ability of the general transcription machinery to form a stable preinitiation complex on the promoter, deletions in this region can be difficult to evaluate. Therefore, it was determined that specific substitution mutations would be more informative. Nevertheless, 3' deletions would have provided additional information for this study and could be beneficial for other studies.

Substitution Mutant Analysis

Specific substitution mutations were introduced into the –150 to +55 regions to identify important control elements (Figs 2 and 3). Although the –150 to +55 fragment retained only 40% of wild-type activity, the activity remained strongly inducible (not shown) and therefore was likely to contain most, if not all, of the key promoter elements involved in inducible transcription. Most of the mutations introduced between –150 and +55 altered 6 bp, although some altered 5 bp and others 10 bp (Fig. 3).

FIGURE 2. Strategy for generating substitution mutants from deletion mutant series.

To construct most of the mutant plasmids, the 5' deletion constructs described above were used as starting points. To generate a substitution mutant from a deletion mutant, PCR was used to amplify the distal portion of the promoter, which was then fused to the appropriate deletion mutant containing the proximal portion. The plasmid generated from this fusion contained a substitution mutation at the site of the fusion. For example, substitution mutant -99/-94s was generated from deletion mutant -93 (Fig. 2). This deletion mutant contains a *PstI* site immediately upstream of nucleotide -93 of the promoter, with a *SacI* site immediately upstream of the *PstI* site. To generate the substitution mutant, the promoter sequence extending from -350 to -100 was amplified by PCR from a full-length promoter template, using an upstream primer containing a *SacI* restriction site and a downstream primer containing a *PstI* site. The PCR product was then inserted into the *SacI*/*PstI*-cleaved -93 deletion mutant plasmid. The *PstI* site generated a 6-bp mutation from -99 to -94 in the context of the -350/+55 promoter fragment. Thus, for this analysis, the deletion mutants served as cloning intermediates for many of the substitution mutants. Alternative strategies for generating substitution mutants (e.g., the Stratagene/Agilent QuikChange method) have been equally successful.

Substitution mutations of 5–10 bp were used for two reasons. First, the mutations needed to be sufficiently small so that important control elements could be localized with reasonable precision. However, it was desirable for them to be sufficiently large so that an unreasonable number of

FIGURE 3. Substitution mutant analysis of the murine *Il12b* promoter. (A) Location and sequence of the substitution mutants. (B) The substitution mutants and their effect, expressed as % of wild-type promoter activity. (Adapted, with permission, from Plevy et al. 1997, ©American Society for Microbiology.)

mutants would not be needed. By using twenty-one 5–10-bp mutants, almost the entire region from –150 to +55 could be scanned for functional elements.

A few gaps are apparent in the mutant series shown in Figure 3. Gaps of 4 bp or less are unlikely to be significant because most sequence-specific DNA-binding proteins recognize sequences of 6 bp or more. Thus, the mutations flanking these small gaps should provide information regarding the existence of a control element. Gaps of 5 bp or more increase the probability that an important control element could be missed. The large gaps remaining after the initial *Il12b* analysis were not intended and were subsequently eliminated by analysis of additional mutations (J. Gemberling and S. Plevy, pers. comm.).

By analyzing the substitution mutants in a transient transfection assay, important control elements were identified (Fig. 3). The most important element for promoter activity in this assay was apparently disrupted by the –99/–94s and –93/–88s mutations. These mutations reduced promoter activity to ~10% of wild type. Another mutation, –132/–127s, reduced promoter activity to 25% of wild type, suggesting that it disrupted another important element. Two other mutations, –107/–102s and –29/–24s, reduced activity to ~25% of wild type. The latter mutation disrupted the TATA box and the former was immediately adjacent to the two severe mutations, suggesting that it might affect the same element.

Many of the remaining mutants showed promoter activities between 50 and 150% of wild type. These small effects suggest the existence of elements that are less important for activity, or elements whose activities are largely redundant (or at least nonsynergistic) with the activities of other elements (see below). Alternatively, as discussed above for the deletion mutants between –350 and –150, these small effects might be caused by the introduction of a foreign sequence into a specific site in the promoter. In other words, the sequences mutated might not contain an important element. Notably, subsequent studies revealed the existence of important binding sites for activator protein-1 (AP-1), nuclear factor of activated T-cells (NFAT), and interferon regulatory factor (IRF) family members (e.g., ICSBP), all located between –88 and the TATA box (Zhu et al. 2001, 2003 [see Fig. 7]).

These elements were missed in the initial analysis because of redundancy, because the substitution mutation originally introduced into one of the elements did not disrupt protein binding to a sufficient extent, and/or because the importance of these elements for *IL12b* promoter activity is sensitive to transfection and cell activation conditions.

Refined Substitution Mutant Analysis

The results in Figure 3 suggest that a critical control element exists between approximately –99 and –88 and that another may exist between approximately –132 and –127. At this point, it could be argued that the mutant analysis has been completed, and experiments should next be performed to identify the proteins that bind the –99/–88 and –132/–127 elements. This argument has some validity. However, it is important to explain why the construction and analysis of additional mutants can provide new and valuable information.

One problem with the results obtained with the 5- to 10-bp substitution mutants is that they provide fairly imprecise information about the boundaries of the important elements. For example, the important nucleotides within the –99/–88 element might extend from –101 to –88, or even further because the –107/–102s mutation reduced promoter activity by fourfold. In addition, the important nucleotides within the –132/–127 element might extend from –135 to –122, because the flanking mutations that had no significant effect on promoter activity were the –121/–117s and –141/–136s mutations.

Before explaining the reason for defining the boundaries of the elements, we describe the strategy and results obtained for the –99/–88 element. A series of mutants was generated, each of which altered three adjacent base pairs (Fig. 4). Analysis of the 3-bp mutant series revealed that three of the mutants showed strongly reduced promoter activities. In contrast, the flanking mutations had no significant effect. These findings suggest that the critical nucleotides span a minimum of 5 bp (i.e., –94 to –90) and a maximum of 9 bp (–96 to –88). This is most consistent with the existence of a binding-site for one protein. In fact, a TRANSFAC database search revealed that the critical 9 bp represent a binding site for CCAAT/enhancer-binding protein (C/EBP) family members (Wedel and Ziegler-Heitbrock 1995). Subsequent DNA-binding studies supported the hypothesis that C/EBP proteins functionally interact with the critical element (Plevy et al. 1997; Bradley et al. 2003). At –130, the functionally important nucleotides identified using a series of 3-bp mutations did not match consensus-binding sites for any known proteins (not shown). However, a closer examination of the specific nucleotide sequence suggested that it represents a nonconsensus-binding site for NF- κ B proteins (Murphy et al. 1995; Plevy et al. 1997; Sanjabi et al. 2005).

FIGURE 4. Refined substitution mutant analysis of the murine *IL12b* p40 promoter. (Adapted, with permission, from Plevy et al. 1997, ©American Society for Microbiology.)

We now return to the question: Why is it worthwhile localizing with precision the boundaries of each element? The answer is that knowledge of the boundaries will help to establish whether each element interacts with one key sequence-specific DNA-binding protein or represents a composite element containing adjacent sites for two or three proteins. This is important because composite elements have been found in promoters and distant control regions with considerable frequency. Composite elements often contain adjacent binding sites for two or more proteins, which bind cooperatively and/or function in a synergistic manner. If an important element is found to span 5–10 bp, the element most likely interacts with one protein (or is a composite element containing coinciding binding sites for two proteins, such as the NF- κ B/HMGA composite sites) (Thanos and Maniatis 1992). On the other hand, if the element spans 15–20 bp or more, it is more likely to be a composite element. Insight into this issue can be gained by analysis of a few extra mutations, making the effort worthwhile. Without this information, one would need to rely on binding studies for insight into the number of proteins that functionally interact with each element. Given the challenge of determining which binding proteins are relevant for a control element (see above), relying solely on binding studies to determine whether a composite element exists is discouraged. In the case of the –99/–88 analysis, the 3-bp mutant results suggested that the element comprised a maximum of 9 bp. When that 9 bp was found to represent a near-consensus binding site for C/EBP proteins, the protein studies could be pursued with considerable confidence. If the boundaries of the element had not been determined, uncertainty would have remained regarding the possibility that the element contained binding sites for additional proteins.

An additional benefit of the 3-bp mutants is worth noting: When binding activities are subsequently identified, the ability of the protein to bind the different mutants can be assessed. A close correlation between the nucleotides required for protein binding and those required for the function of the element in a transfection experiment provides an important piece of data that can support the functional relevance of the protein. In this case, the binding of recombinant C/EBP proteins required precisely the same nucleotides as were required for promoter function, supporting the hypothesis that C/EBP proteins are responsible for the function of the element.

One final issue should be discussed with regard to the refined mutant analysis: Why use 3-bp mutants rather than single-base-pair mutants? Using single-base-pair substitutions would require three times as many mutants to scan the important region. Admittedly, the boundaries of the element would be defined more precisely with single-base-pair mutants, but for many studies, the benefit might not outweigh the additional effort that would be needed. In addition, many proteins can tolerate single-base-pair changes at some positions within their binding sites with only a minor loss of activity. Furthermore, if single-base-pair changes were used, the results would have been dependent on the particular nucleotide introduced, because binding proteins often tolerate some substitutions better than others at a given position. The probability that a given 3-bp mutation will be tolerated by a DNA-binding protein is much lower.

Choice of Nucleotides for Substitution Mutants

Unfortunately, there is no foolproof strategy for nucleotide choice when constructing substitution mutants. The possibility will always exist that a mutation will create a fortuitous binding site for another protein that might influence promoter activity. The creation of a new binding site could lead to inaccurate or misleading data. In the 5–10-bp substitution mutant series, a restriction enzyme site was inserted in place of the *Il12b* promoter sequences. In this case, the restriction site was necessary for mutagenesis, because the substitution mutants were generated from deletion mutants using a technique that relied on the presence of a restriction site (see Fig. 2). For most mutagenesis approaches, insertion of a restriction site is convenient because it simplifies the process of screening for bacterial colonies containing mutant plasmids; minipreps of the DNAs can simply be analyzed by restriction mapping, rather than DNA sequencing.

For the 3-bp mutants, restriction sites could not be introduced routinely because only 3 bp were altered. Because of the small size of the substitution, it also was important to alter the sequence as

severely as possible, to increase the probability that important protein–DNA contacts would be disrupted. To this end, each substitution introduced a transversion and, at the same time, changed the base pair. For example, a C/G base pair was changed to A/T, G/C to T/A, A/T to C/G, and T/A to G/C. These substitutions introduce the most radical structural change possible at a particular site. However, this does not take into account the possibility that the substitution will result in the creation of a binding site for another protein.

To determine whether a binding site for another protein is generated, the mutant sequence can be analyzed by searching a binding-site database such as TRANSFAC. If the mutant sequence is similar to the consensus sequence for a known protein, a different sequence should be introduced. Of course, this analysis will only reveal binding sites for known proteins, based on current knowledge. Because a definitive method is not available for ensuring that a substitution mutation does not introduce another binding site, it is important to keep this possibility in mind during the subsequent steps in the promoter analysis. If a result at any stage of the analysis suggests that a fortuitous binding site might have been introduced accidentally, the best course of action is to address this possibility by preparing additional substitution mutations at the same location.

Inducibility and Cell-type Specificity

The mutagenesis strategy outlined above resulted in the identification of two control elements that contribute to *Il12b* promoter activity in LPS-induced macrophages. The promoter is inducible; thus, both of these elements may bind proteins that directly mediate induction. Alternatively, one of the elements may bind a constitutively active protein that is essential for promoter function but is not a direct mediator of induction. With this latter scenario, promoter induction would occur when the constitutively active protein carries out an appropriate physical or functional interaction with the induced protein. Because the primary objective of the promoter analysis is to elucidate the mechanism of promoter induction, it eventually will be necessary to distinguish between these two types of elements. Similar issues must be considered when studying cell type-specific control regions. In these studies, the goal is to distinguish the control elements that bind cell type-specific proteins from those that bind ubiquitously active proteins. It is important to add that some factors can contribute to both transcriptional repression and transcriptional activation of the same gene through the same DNA motif, further increasing the challenge of understanding the precise function of each motif.

It can be difficult to determine which elements mediate induction directly because the activity of a control region following induction is sensitive to mutations in any control element, including those that do not mediate induction but are merely required for activity. Because the importance of an element following induction provides no significant information regarding inducibility per se, a common strategy for determining which elements are direct mediators of induction is to rely on the importance of each element before induction. In theory, elements that bind proteins that mediate induction directly will not be involved in basal promoter activity in uninduced cells and will become important only following induction. In contrast, elements that bind constitutively active proteins will be equally important in uninduced and induced cells. Thus, it is thought that the precise role of each element can be determined simply by comparing the “fold-induction” of the wild-type promoter to that of each promoter mutant. In other words, after subtracting background, the induced reporter activity of each construct is divided by the uninduced activity of the same construct, yielding a fold-induction value. If mutation of an element reduces the fold-induction value (i.e., if the element is more important in induced cells than uninduced cells), the element is considered to be a direct mediator of induction. If mutation of an element does not influence the fold-induction value (i.e., if the element is equally important in uninduced and induced cells), the element is unlikely to mediate induction.

The above strategy can be informative for the subset of promoters that yield considerable activity before induction, but it is not useful for promoters whose uninduced activities are not substantially above background. The activity of the *Il12b* promoter, for example, was only slightly greater than background before induction. Although the uninduced activity of the wild-type promoter was

statistically significant, mutations in the important control elements reduced the uninduced signal to levels that were not significantly greater than background. This scenario, which is quite common, renders an analysis of the fold-induction values meaningless.

As a specific example, consider the mutation in the NF- κ B (Rel) site. This mutation reduces the induced promoter activity to 25% of wild type (see Fig. 5). It also reduces the uninduced signal to a value only slightly greater than background. Because the uninduced signal for the promoter mutant is nearly zero after subtracting background, the fold-induction remains high. If those values were used as the sole criterion for determining whether an element mediates inducibility, one would conclude that the NF- κ B site was not important for induction, despite considerable evidence that NF- κ B proteins play critical, direct roles during gene induction in macrophages.

The results obtained with a mutation in the C/EBP site provide another example. With this mutation, induced promoter activity is reduced to <10% of wild type (Fig. 5). Uninduced promoter activity is also reduced, but like the NF- κ B mutation, it remains slightly above background. The fold-induction calculations yield a value for the C/EBP mutant that is much lower than for the NF- κ B mutant. One interpretation of these data is that the C/EBP site is the key to induction, with the NF- κ B site much less important. Although this seems logical, it actually represents a misinterpretation of the data, because the fold-induction values depend on the precision of the uninduced signals. Because the uninduced signals are close to background with both the NF- κ B and C/EBP mutants, as well as with the wild-type promoter, the accuracy of these numbers is difficult to determine, even with statistical analyses. In other words, very small changes in the uninduced signals can have dramatic, but highly questionable, effects on fold-induction values.

Regardless of the results obtained, the activities of promoter mutants rarely provide substantive insight into the issue of which elements are directly responsible for inducibility (or cell-type specificity). To address this issue, the relevant binding proteins must be identified and their properties characterized. If the abundance of the relevant transcription factor increases during cell induction, it may contribute directly to promoter induction. This hypothesis can be tested more rigorously using other approaches, which are not discussed here. If transcription factor abundance is not increased during cell induction, the factor may nevertheless play a direct role in induction because it may acquire a posttranslational modification that alters its activity. A careful analysis of each binding site and transcription factor is ultimately needed to determine which ones play direct roles in induction or cell-type specificity.

FIGURE 5. Effect of *I12b* mutations on promoter induction by lipopolysaccharide.

Transcription Start-Site Confirmation

Ideally, the start sites of transcripts should be determined when analyzing a promoter by transient transfection, to confirm that transcription initiates at the correct location. The transcription start sites of mutant promoters should also be determined to confirm that the mutations do not alter the start site or result in the induction of a cryptic start site. Minor start-site alterations on promoter mutagenesis are of little concern, but more severe changes may indicate that a mutation has not simply disrupted a control element. Instead, the mutation may have altered the overall structure and regulation of the promoter. For example, a mutation might lead to the activation of a cryptic TATA-like sequence within the mutant nucleotides or elsewhere in the promoter. If a cryptic TATA-like sequence becomes activated, it may respond to the regulatory elements differently from the authentic core promoter.

Although the transient transfection efficiencies of many cell lines are too low to allow start-site mapping by primer extension or RNase protection, it should be possible to use 5' RACE (rapid amplification of cDNA ends) to gain insight into the start sites being used. As an alternative to the sensitive 5' RACE approach, stable transfectants can be prepared with a promoter–reporter plasmid. Because every cell in the selected lines will contain an integrated reporter plasmid, the reporter transcripts should be of sufficient abundance for start-site mapping by primer extension or RNase protection. This strategy was used for the *Il12b* promoter analysis, and the primer extension results confirmed that the major start site was at the expected location. The results provided some confidence that the transcription start sites in the transiently transfected plasmids were probably also at the correct location.

Because the locations of transcription start sites can be difficult to determine following transient transfection, this experiment is usually not performed during a typical mutant analysis. For studies that involve the dissection of a core promoter region (i.e., TATA and Inr region), the absence of information regarding the start-site location is likely to be problematic. However, for most other studies, the start-site analysis is not essential. Nevertheless, one should proceed with considerable caution and remain aware of the fact that the start site has not been confirmed.

Choice of Assay

One key limitation of the mutagenesis strategy is that the only elements identified are those that are important in the functional assay being used. In the *Il12b* promoter analysis, which used a transient transfection assay, the C/EBP site was essential, whereas the NF- κ B site made only a moderate contribution. In contrast, when mutations in these two sites were tested in a stable transfection assay, both were absolutely essential for promoter activity (Plevy et al. 1997; Sanjabi et al. 2005). Presumably, the high plasmid copy number in transiently transfected cells, or the episomal nature of the transiently transfected plasmids, diminished the importance of the NF- κ B site. It would not be surprising if repetition of the entire mutant analysis using a stable transfection assay resulted in the identification of other essential control elements, which were relatively unimportant in the transient assay. Indeed, several examples of control elements that function in stable, but not transient, transfection assays have been reported.

Some control elements that are important for transcription of the endogenous gene may be missed in both transient and stable transfection assays. This is because stably transfected plasmids do not become incorporated into the same chromatin structure as the endogenous gene. Therefore, a subset of the control elements important for chromatin remodeling during gene activation or inactivation might be missed.

One example of the above is provided by an analysis of the immunoglobulin (Ig) μ intronic enhancer. In transgenic mice, the activity of the Ig μ enhancer was strongly stimulated by the adjacent matrix attachment regions (MARs) (Forrester et al. 1994, 1999; Fernandez et al. 1998). However, in both transient and stable transfection assays, the MARs had no effect on enhancer activity. Thus, if the Ig μ enhancer had been dissected solely with transfection assays, the importance of the MARs would have been missed. Interestingly, Forrester and Grosschedl developed a modified stable

transfection assay that restored the MAR requirement (Fernandez et al. 1998; Forrester et al. 1999). For this assay, the reporter plasmid DNA was methylated *in vitro* before transfection and drug selection. Reporter gene activity from the stably integrated, premethylated plasmids required the presence of the MARs. Apparently, premethylation caused the transfected plasmid to become incorporated into less accessible chromatin, which resulted in the MAR requirement for transcriptional activation.

On the basis of the above results, a mutant analysis of a new control region would ideally be performed with a transgenic mouse assay, or at least a stable transfection assay. Unfortunately, the time and resources required for a comprehensive mutant analysis by either of these assays make them impractical for most studies. For this reason, it often is necessary to begin with a transient transfection assay to identify the control elements needed in that assay and then to proceed to more sophisticated assays when the analysis reaches a more advanced stage.

Redundancy of Control Elements

A final caveat of the comprehensive mutant analysis is that it may fail to identify control elements whose activities are redundant (or at least are not strongly synergistic) with the activities of other control elements within the region. The Ig μ enhancer provides a classic example of redundancy within a control region. Early transfection studies revealed that no substitution mutations reduced activity by more than approximately twofold (e.g., see Lenardo et al. 1987). Similar results were obtained in transgenic mouse assays (Jenuwein and Grosschedl 1991; Annweiler et al. 1992), suggesting that the apparent redundancy was not an experimental artifact of the transfection assays. The inability of any mutation to strongly diminish enhancer activity created considerable difficulties for the analysis of the mechanism of Ig μ enhancer function.

Sen and colleagues, however, pursued a strategy for circumventing the redundancy problem and for dissecting the molecular basis of the redundancy (Fig. 6). They first created deletion mutants to identify the smallest enhancer fragment that supports enhancer function preferentially in B cells (Nelsen et al. 1990, 1993). As expected on the basis of the observed redundancy, several control elements could be deleted with little consequence. Substitution mutations were then introduced into the minimal enhancer fragment, revealing that the remaining control elements were absolutely essential for enhancer function (Nelsen et al. 1990, 1993). Further analysis of the essential control elements identified proteins that may functionally interact with them (Nelsen et al. 1990, 1993). The molecular mechanism by which these proteins synergize with one another has also been dissected using the minimal enhancer fragment (Erman and Sen 1996; Nikolajczyk et al. 1996, 1997; Rao et al. 1997; Erman et al. 1998). Initial studies used a minimal enhancer fragment that contained only three control elements: μ A, μ E3, and μ B (Fig. 6). With this small fragment, it was necessary to fuse dimers to the reporter plasmid to detect activity. After identifying and characterizing the three elements within this fragment, a larger fragment was used, which yielded substantial enhancer activity when present in a single copy (Dang et al. 1998b). Because this larger fragment still lacked the redundant elements, most of the elements remained essential for enhancer activity.

To understand the molecular basis of redundancy, Sen and colleagues mutated the full-length enhancer systematically to identify the control element that is redundant with μ E3 (Fig. 6). The goal was to identify the elements critical for function when the enhancer contained a mutant μ E3 site. The element that conferred redundancy was a previously undescribed enhancer element that binds IRF proteins (Dang et al. 1998a). Because the IRF element is largely redundant with the μ E3 element, it appears to be just as important for enhancer function, even though it was discovered 13 years after the μ E elements were first reported.

It is noteworthy that the μ E elements were originally identified by *in vivo* and *in vitro* protein–DNA interaction studies. In contrast, discovery of the IRF element required a systematic mutant analysis that was sufficiently comprehensive to address the redundancy issue. This systematic analysis can now be extended to determine whether other elements contribute to redundancy within the enhancer. The biological basis for the redundancy remains unknown. One hypothesis is that it allows the enhancer to be activated by distinct combinations of factors at different stages of

FIGURE 6. Strategy for dissecting an enhancer that shows considerable redundancy among control elements.

development, so that the same specific set of factors does not need to be present whenever the enhancer is activated. Alternatively, redundancy may ensure enhancer function in a nuclear milieu of limiting transcription factor concentrations.

The strategy used by Sen and colleagues is likely to be useful for analyzing redundant (or non-synergistic) elements in other control regions. Some redundancies may be biologically significant, whereas others may be related to the assay used for the analysis. The Ig μ enhancer appears to be an example of biologically relevant redundancy, because the redundancy was observed in transgenic mouse assays, as well as in transient and stable transfection assays.

FIGURE 7. Information provided by a comprehensive mutant analysis of the *Il12b* promoter versus a phylogenetic analysis using the University of California, Santa Cruz Genome Browser (<http://genome.ucsc.edu>; Kent et al. 2002).

INSIGHTS PROVIDED BY A COMPREHENSIVE MUTANT ANALYSIS VERSUS A PHYLOGENETIC ANALYSIS

We conclude this article by addressing the level of success of the comprehensive mutant analysis of the *Il12b* promoter described above in comparison with information gained from a phylogenetic analysis and binding site database analysis. Figure 7 shows a phylogenetic analysis of the *Il12b* promoter generated by the University of California, Santa Cruz Genome Browser. The locations of DNA elements originally identified in the comprehensive mutant studies described in this article and in other published studies are depicted above the sequence (Zhu et al. 2001, 2003; Sanjabi et al. 2005). Interestingly, only the TATA box and C/EBP site were readily detected in a basic TRANSFAC database search; the other sites differed from the reported consensus sequences for proteins now known to be capable of binding the sites. A few additional sites were identified in TRANSFAC analyses performed with reduced stringencies (not shown), but in the reduced stringency analyses, these functionally important sites did not stand out in relation to predicted nonconsensus sites for a large number of other transcription factors. As one example, the TRANSFAC analysis predicted NF- κ B-binding sites at locations that are not conserved through evolution and that were not important for promoter activity in transient transfection assays, but it missed the functionally important sites that can bind NF- κ B.

Significantly, the phylogenetic analysis revealed that all of the functionally important transcription factor binding sites have been highly conserved through mammalian evolution. Therefore, the phylogenetic analysis would have successfully predicted that these elements are functionally important if the genome sequences for such a large number of species had been available before the mutant studies were performed. Interestingly, the phylogenetic analysis reveals several additional DNA motifs that are just as highly conserved (horizontal lines numbered 1 through 8 in Fig. 7), even though disruption of these motifs by mutagenesis had no significant effect on promoter activity in a transient transfection assay (see Figs 1 and 3) (Plevy et al. 1997).

The above observations raise a critical question: Are the highly conserved elements that were not important for *Il12b* promoter activity in a transient transfection assay ever important for promoter activity? One possibility is that these elements contribute to promoter activity in a different cell type or in response to a different stimulus. Alternatively, they may be important for transcription only in a more native chromosomal environment, in which they may facilitate interactions between the promoter and distant enhancers, or they may contribute to nucleosome remodeling events that are not necessary for transcription in a transfection assay. A final possibility is that these elements may indeed be important in a transfection assay (and also at the endogenous locus), but their functions may not be apparent because they function redundantly with other promoter elements. To distinguish between these possibilities, it will be necessary to carefully evaluate the functions of these highly conserved elements in other situations, in particular through mutagenesis in a native chromosomal environment.

On the basis of this comparison, we conclude that elucidating the mechanism regulating any mammalian gene requires a variety of experimental approaches. Initial insight can be gained by performing phylogenetic analyses, binding-site database analyses, and ChIP experiments to evaluate the predicted protein–DNA interactions. A comprehensive mutant analysis using a transfection assay is likely to reveal at least a few key elements that play particularly important roles in transcriptional regulation and can allow the researcher to initiate, with confidence, more advanced studies of the factors that bind these elements. However, to fully understand the regulation of a gene, strategies must be developed for exploring the functional roles of the many other DNA motifs that are conserved through evolution but that do not contribute significant functions in standard transfection assays. The first step toward this goal will likely be to perform comprehensive mutant studies of conserved elements in a native chromosomal environment.

REFERENCES

Annweiler A, Muller U, Wirth T. 1992. Functional analysis of defined mutations in the immunoglobulin heavy-chain enhancer in transgenic mice. <i>Nucleic Acids Res</i> 20: 1503–1509.	Bradley MN, Zhou L, Smale S.T. 2003. C/EBPβ regulation in lipopolysaccharide-stimulated macrophages. <i>Mol Cell Biol</i> 23: 4841–4858.

M.F. Carey et al.

- Carey MF, Peterson CL, Smale ST. 2012a. Experimental strategies for the identification of DNA-binding proteins. *Cold Spring Harb Protoc* doi: 10.1101/pdb.top067470.
- Carey MF, Peterson CL, Smale ST. 2012b. Experimental strategies for cloning or identifying genes encoding DNA-binding proteins. *Cold Spring Harb Protoc* doi: 10.1101/pdb.top067900.
- Dang W, Nikolajczyk BS, Sen R. 1998a. Exploring functional redundancy in the immunoglobulin m heavy-chain gene enhancer. *Mol Cell Biol* 18: 6870–6878.
- Dang W, Sun XH, Sen R. 1998b. ETS-mediated cooperation between basic helix-loop-helix motifs of the immunoglobulin m heavy-chain gene enhancer. *Mol Cell Biol* 18: 1477–1488.
- Erman B, Sen R. 1996. Context dependent transactivation domains activate the immunoglobulin m heavy chain gene enhancer. *EMBO J* 15: 4565–4575.
- Erman B, Cortes M, Nikolajczyk BS, Speck NA, Sen R. 1998. ETS-core binding factor: A common composite motif in antigen receptor gene enhancers. *Mol Cell Biol* 18: 1322–1330.
- Fernandez LA, Winkler M, Forrester W, Jenuwein T, Grosschedl R. 1998. Nuclear matrix attachment regions confer long-range function upon the immunoglobulin m enhancer. *Cold Spring Harbor Symp Quant Biol* 63: 151–524.
- Forrester WC, van Genderen C, Jenuwein T, Grosschedl R. 1994. Dependence of enhancer-mediated transcription of the immunoglobulin m gene on nuclear matrix attachment regions. *Science* 265: 1221–1225.
- Forrester WC, Fernandez LA, Grosschedl R. 1999. Nuclear matrix attachment regions antagonize methylation-dependent repression of long-range enhancer–promoter interactions. *Genes Dev* 13: 3003–3014.
- Jenuwein T, Grosschedl R. 1991. Complex pattern of immunoglobulin m gene expression in normal and transgenic mice: Nonoverlapping regulatory sequences govern distinct tissue specificities. *Genes Dev* 5: 932–943.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D. 2002. The Human Genome Browser at UCSC. *Genome Res* 12: 996–1006.
- Lenardo M, Pierce JW, Baltimore D. 1987. Protein-binding sites in Ig gene enhancers determine transcriptional activity and inducibility. *Science* 236: 1573–1577.
- Ma X, Chow JM, Gri G, Carra G, Gerosa F, Wolf SF, Dzialo R, Trinchieri G. 1996. The interleukin 12 p40 gene promoter is primed by interferon g in monocytic cells. *J Exp Med* 183: 147–157.
- Murphy TL, Cleveland MG, Kulesza P, Magram J, Murphy KM. 1995. Regulation of interleukin 12 p40 expression through an NF- κ B half-site. *Mol Cell Biol* 15: 5258–5267.
- Nelsen B, Kadesch T, Sen R. 1990. Complex regulation of the immunoglobulin μ heavy-chain gene enhancer: μ B, a new determinant of enhancer function. *Mol Cell Biol* 10: 3145–3154.
- Nelsen B, Tian G, Erman B, Gregoire J, Maki R, Graves B, Sen R. 1993. Regulation of lymphoid-specific immunoglobulin μ heavy chain gene enhancer by ETS-domain proteins. *Science* 261: 82–86.
- Nikolajczyk BS, Nelsen B, Sen R. 1996. Precise alignment of sites required for m enhancer activation in B cells. *Mol Cell Biol* 16: 4544–4554.
- Nikolajczyk BS, Cortes M, Feinman R, Sen R. 1997. Combinatorial determinants of tissue-specific transcription in B cells and macrophages. *Mol Cell Biol* 17: 3527–3535.
- Plevy SE, Gemberling JH, Hsu S, Dorner AJ, Smale ST. 1997. Multiple control elements mediate activation of the murine and human interleukin 12 p40 promoters: Evidence of functional synergy between C/EBP and Rel proteins. *Mol Cell Biol* 17: 4572–4588.
- Rao E, Dang W, Tian G, Sen R. 1997. A three-protein-DNA complex on a B cell-specific domain of the immunoglobulin μ heavy chain gene enhancer. *J Biol Chem* 272: 6722–6732.
- Sanjabi S, Williams KJ, Sacconi S, Zhou L, Hoffmann A, Gerondakis S, Natoli G, Smale ST. 2005. A c-Rel subdomain responsible for enhanced DNA-binding affinity and selective gene activation. *Genes Dev* 19: 2138–2151.
- Thanos D, Maniatis T. 1992. The high mobility group protein HMG I(Y) is required for NF- κ B-dependent virus induction of the human IFN- β gene. *Cell* 71: 777–789.
- Wedel A, Ziegler-Heitbrock HW. 1995. The C/EBP family of transcription factors. *Immunobiology* 193: 171–185.
- Zhu C, Gagnidze K, Gemberling JH, Plevy SE. 2001. Characterization of an activation protein-1-binding site in the murine interleukin-12 p40 promoter. Demonstration of novel functional elements by a reductionist approach. *J Biol Chem* 276: 18519–18528.
- Zhu C, Rao K, Xiong H, Gagnidze K, Li F, Horvath C, Plevy S. 2003. Activation of the murine interleukin-12 p40 promoter by functional interactions between NFAT and ICSBP. *J Biol Chem* 278: 39372–39382.

Cold Spring Harbor Protocols

Identifying *cis*-Acting DNA Elements within a Control Region

Michael F. Carey, Craig L. Peterson and Stephen T. Smale

Cold Spring Harb Protoc; doi: 10.1101/pdb.top068171

Email Alerting Service

Receive free email alerts when new articles cite this article - [click here](#).

Subject Categories

Browse articles on similar topics from *Cold Spring Harbor Protocols*.

[Analysis of Gene Expression in Cultured Cells](#) (66 articles)
[DNA:Protein Interactions](#) (74 articles)
[Gene Fusion](#) (17 articles)
[Genome Analysis](#) (191 articles)
[Mutagenesis](#) (57 articles)
[Mutagenesis by PCR](#) (20 articles)
[Sequence Database Searching](#) (38 articles)
[Use of Reporter Genes](#) (124 articles)

To subscribe to *Cold Spring Harbor Protocols* go to:
<http://cshprotocols.cshlp.org/subscriptions>